LOCALIZATION OF PARAMAGNETIC METAL IONS IN BIOMOLECULES BY MEANS OF ELECTRON PARAMAGNETIC RESONANCE

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

DINAR ABDULLIN

aus

Zelenodolsk, Russland

Bonn 2017

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter	Prof. Dr. Olav Schiemann
2. Gutachter	Prof. Dr. Thomas Prisner
Tag der Promotion:	3. Februar 2017
Erscheinungsjahr:	2017

In memory of my grandfather

Abstract

Localization of metal ion binding sites in biomolecules is an important objective in structural biology. Experimental methods that are commonly used to determine the location of metal ions, such as X-ray crystallography and NMR, are not always applicable due to their limitations related to sample preparation and data analysis. It is, therefore, of high importance to have complementary methods that can be applied to difficult cases. This thesis reports on the first application of EPR spectroscopy for the localization of metal ions in biomolecules via trilateration. Since the trilateration principle implies the use of distance constraints to determine the location of a metal ion, a significant part of the work is related to distance measurements between metal ions and spin labels by means of pulsed EPR techniques. The most popular technique for such measurements, pulsed electron-electron double resonance (PELDOR), is often affected by orientation selectivity effects which significantly complicate the derivation of distances from the experimental time-domain data. To enable the extraction of distances from orientation-selective PELDOR data without any preliminary knowledge about the relative orientation of spin centers, a method based on a simplified geometric model of a spin system was developed. The method was implemented in a computer program called PeldorFit, and its validity was confirmed by multiple tests on PELDOR data sets from the literature and from the present work. In particular, the new method allowed extracting Cu^{2+}-MTSSL distance distributions from PELDOR data sets of six MTSSL-labeled azurin mutants. These distance distributions were then used for benchmarking the EPR-based trilateration on the Cu^{2+} ion of azurin. The trilateration problem was solved with the newly developed program mtsslTrilaterate. The obtained Cu^{2+} coordinates were compared to the available crystal structure of azurin, revealing that the calculated Cu^{2+} location is very close to the corresponding crystallographic site. The precision of the trilateration was estimated at 0.26 nm. This value was shown to be affected by the spin density delocalization of the Cu^{2+} center, the precision of the average MTSSL coordinates, the number and precision of the distance constrains, and the accuracy of the structural model of the metal-free protein. The main source of the trilateration error was, however, attributed to the average MTSSL coordinates. These coordinates were determined using the existing *in silico* spin labeling methods which were shown to have an average error of 0.15 nm. More accurate MTSSL coordinates were obtained using the PELDOR constraints and measuring the crystal structures of the azurin mutants. Furthermore, the possibility to extend the EPR-based trilateration to other metal ions, such as Fe^{3+}, was investigated in the context of the distance measurements

for these ions. Since PELDOR distance measurements on the Fe^{3+}/nitroxide spin pair are complicated due to the large spectral width of Fe^{3+}, the use of an alternative pulsed EPR technique RIDME was considered. Both techniques, RIDME and PELDOR, were compared using a MTSSL-labeled mutant of the heme-containing protein cytochrome P450cam. The results of the comparison are that RIDME allows avoiding the orientation selectivity effects and yields a seven times higher signal-to-noise ratio than PELDOR. This demonstrates the great potential of the RIDME technique for future trilateration studies. Besides distance constraints, the availability of spin labels with well-studied structure and dynamics is required for the trilateration. Since one of the potential applications of the EPR-based trilateration is the localization of metal centers in ribozymes, the structure and dynamics of a nitroxide-labeled uracil nucleobase **dU** were investigated in this work. Using the information obtained by EPR spectroscopy and DFT calculations, a structural model of **dU** was proposed and implemented in the program mtsslWizard. This model was successfully applied for the prediction of the mean distance between two **dU** nucleobases in an RNA duplex. The comparison of the predicted distance with the corresponding PELDOR-derived distance revealed only a minor difference of 0.1 nm. Thus, the performed characterization of the **dU** nucleobase makes it suitable for future EPR-based metal ion localization in ribozymes.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Olav Schiemann for his great support and guidance during my studies.

A significant part of this work was done with great help from my co-workers. It gives me pleasure to thank Dr. Gregor Hagelüken for taking care about all biosynthesis, as well as for always exciting discussions. Next, I would like to thank Nicole Florin and Fraser Duthie, who prepared all protein samples which were investigated in this work. Special thanks go to Dr. Hideto Matsuoka and Dr. Yaser NejatyJahromy for maintaining the EPR spectrometers and for helpful technical advice. Finally, I would like to express my deep appreciation for all members of the group for the nice work atmosphere.

Further, I would like to thank all collaborators from other research groups: Prof. Dr. Robert Glaum and Dr. Subrata Roy for the work on vanadyl phosphates, Prof. Dr. Graham Smith and Dr. Robert Hunter for showing me their outstanding EPR spectrometer and providing me the data measured on it, Prof. Dr. Arne Lützen and Christoph Klein for the work on Cu(II) helicates and Fe(III) porphyrins, Prof. Dr. Michael Famulok and Mark Kerzhner for the work on a new spin label for RNA.

Special recognition is given to the University of Bonn for giving me the opportunity to carry out my research and to Deutsche Forschungsgemeinschaft for financing my work within the Collaborative Research Center SFB813 “Chemistry at Spin Centers”.

Last but not least, I would like to express my deepest gratitude to my family and my friends for supporting me spiritually during these years.

Table of contents

1 Introduction

1.1 Motivation

Metal ions are present in many proteins and nucleic acids. They are involved in many essential functions ranging from structure stabilization to enzyme catalysis, signal transduction, muscle contraction, hormone secretion, taste and pain sensation, respiration, and photosynthesis.[1–4] A large part of these functions corresponds to metal ions that bind to biomolecules specifically and tightly (dissociation constants of $< 1\ \mu M$). Such specific binding was observed for a number of different ions as e.g. alkali metals, alkaline earth metals, and transition metals.[1–4] Typically, the tightly bound metal ions feature an interrelation between biological function and their location in the biomolecular structures. Therefore, the investigation of biomolecular processes on a molecular level usually requires knowledge about the metal ion binding sites. The localization of these sites is, however, a challenging task that sometimes requires using a broad arsenal of experimental techniques.

The most commonly used technique for this purpose is X-ray crystallography.[5] The advantage of this technique is that it provides an atomistic resolution of metal ion binding sites and yields the positions of metal ions with high precision. However, this technique requires crystallization of biomolecules, which can be a very demanding task, especially for flexible biomolecules. It is also not always possible to crystallize different conformational states of biomolecules that they adopt during folding or function. Moreover, not all metal ions can be unambiguously identified in the electron density maps. For example, Mg^{2+} ions have a comparable density to that of hydrated Na^{+} ions and water molecules.[6] In these cases, metal-binding sites are usually predicted based on the characteristic coordination of metal ions or by replacing native metal ions with non-native ions which have a higher electron density.[6]

Nuclear magnetic resonance spectroscopy (NMR) is another technique that holds potential for the localization of metal-binding sites. This technique is applicable to aqueous solutions and provides high-resolution structures of biomolecules with atomic masses up to roughly 100 kDa.[7] Importantly, the determination of metal ion binding sites by NMR differs for biomolecules that contain diamagnetic metal ions and biomolecules that contain paramagnetic metal ions. The binding of diamagnetic ions to biomolecules is usually detected through changes of the chemical shifts in NMR spectra.[8] The shortcoming of this method is that the altered NMR spectra may result not only from a specific binding event but also from a global conformational change of a biomolecule. The binding of paramagnetic ions to biomolecules has a more drastic effect on the NMR spectra because these ions affect not only the chemical

shifts but also significantly increase the relaxation rates of the surrounding nuclei.[9] Although both effects can be taken into account in the structure refinement by considering pseudocontact shifts and residual dipolar couplings, the lack of data at very short distances from tightly bound ions often makes it impossible to determine the exact position of the metal site.[9] In addition, both NMR approaches require a reference sample that is either a metal-free form of a biomolecule or a biomolecule in which a paramagnetic metal ion is substituted by a diamagnetic one.

Due to the mentioned limitations of X-ray crystallography and NMR, complementary biochemical and spectroscopic methods for the localization of metal ion binding sites are of high interest. One popular biochemical method is a rescue experiment, in which potential metal-binding sites are examined through changing the ligands of metal ions. The modified ligands disfavor the binding of an intrinsic metal ion, reducing the biological activity of a biomolecule. The activity is restored by adding other metal ions with a sufficiently high binding affinity to the modified ligands. Rescue experiments have been successfully used to identify Mg^{2+} sites in a number of ribozyme systems.[6] However, these experiments are sensitive only to metal ions that contain inner-sphere contacts to modified ligands, which make this method dependent on an initial hypothesis.

Spectroscopic methods, such as fluorescence resonance energy transfer (FRET) and electron paramagnetic resonance (EPR), offer another possibility to localize metal ions through distance measurements between metal ions and fluorescent or paramagnetic labels. The advantage of FRET measurements is that they can be performed in liquid solutions with single molecule sensitivity. Most of the FRET measurements are, however, done on the lanthanide ions which are a steric and chemical analog of the biologically relevant Mg^{2+} and Ca^{2+} ions.[10] Recently, also FRET measurements for Ni^{2+} and Cu^{2+} have been reported.[11,12] In contrast, EPR-based distance measurements are mostly applied to frozen solutions but can be used for a broad range of native metal ions,[13,14] since many transition metal ions are found in the paramagnetic state. Spin labels commonly used in EPR are smaller and more rigid than chromophores, allowing easier correlation of the measured distance with the structure of the studied biomolecule. In addition, the EPR approach can be applied in frozen solution, is not restricted by the biomolecular size and does not require the use of reference samples. Recently, EPR distance constraints were used to localize a nitroxide-labeled lipid in soybean seed lipoxygenase-1 via trilateration[15] and also to narrow down the location of a Cu^{2+} ion in the EcoRI endonuclease-DNA complex.[16] However, there have been no reports on an EPR-based trilateration of a metal ion in a biomolecule and the precision of the trilateration

approach has not been evaluated. Thus, an extension of the EPR-based trilateration approach towards metal ions is of high interest. This would enrich the arsenal of the existing methods and provide to structural biologist an additional tool for studying metal ions in cases when other methods are inapplicable.

1.2 Paramagnetic metal ions in biomolecules

Many of the metal ions that bind to proteins and nucleic acids are transition-metal ions.[1–4] These ions are often found in the paramagnetic state. A list of the most commonly appearing paramagnetic metal ions and examples of biomolecules that contain these ions is given in Table 1.2.1. Note that, except Fe^{2+}, all listed metal ions have an odd number of electrons in their 3d atomic orbital and, therefore, are Kramers ions. Kramers ions have a non-integer spin and their energy levels are at least two-fold degenerated in a zero magnetic field. If a static magnetic field is applied, the degeneracy of these energy levels is lifted and the transitions between the split energy levels can usually be detected by EPR. The metal ions in Table 1.2.1 can be divided into two classes: low-spin ($S = 1/2$) and high-spin ($S > 1/2$). The theoretical description of the EPR experiments on low-spin metal ions is usually simpler than that for high-spin metal ions. For example, the phenomenon of zero-field splitting has to be taken into account only for the high-spin ions (see Section 1.4.1). Note that for some metal ions, such as Fe^{3+} and Co^{2+}, both spin states can be found, in dependence of their coordination.

Additionally, it is sometimes possible to replace intrinsic diamagnetic metal ions, such as Zn^{2+} and Mg^{2+}, by paramagnetic ones, such as Co^{2+} or Mn^{2+}, without drastic change in the biomolecular activity.[6] This possibility enlarges the scope of EPR to a broader range of proteins and oligonucleotides.

In the present work, two low-spin metal ions, the Cu^{2+} ion of the protein azurin and the Fe^{3+} ion of the protein cytochrome P450cam, are considered. Both metalloproteins and their metal ions were characterized previously by a number of spectroscopic methods including EPR. A brief summary of the biological functions and characteristic magnetic properties of these ions is given in the next two sections.

Table 1.2.1. Electronic configurations and spin numbers of biologically relevant paramagnetic metal ions.

Metal ion	Electronic configuration	Spin	Example of a biomolecule
Fe^{3+}	$3d^5$	1/2 or 5/2	Cytochrome P450cam
Fe^{2+}	$3d^6$	2	Cytochrome P450cam
Mn^{2+}	$3d^5$	5/2	Concanavalin A
Co^{2+}	$3d^7$	1/2 or 3/2	Glutamate mutase
Ni^{3+}	$3d^7$	1/2	Hydrogenase
Cu^{2+}	$3d^9$	1/2	Azurin
Mo^{5+}	$3d^1$	1/2	Xanthine oxidase
V^{4+}	$3d^1$	1/2	Vanadium haloperoxidase

1.2.1 The Cu^{2+} ion in azurin

The blue copper protein azurin is important as an electron carrier in biological systems. The electron transfer is accompanied by a change of the valence state of the copper ion between 2+ and 1+. The high resolution X-ray structures of azurin reveal that the copper ion has a distorted trigonal planar coordination with three strong ligands, namely the δ-nitrogen atoms of H46 and H117 and the sulfur of C112; two relatively distant ligands, the carbonyl oxygen of G45 and the sulfur of M121, occupy the axial positions (Figure 1.2.1). The unpaired 3d electron of the Cu^{2+} ion is substantially delocalized onto the sulfur ligand of the residue C112, resulting in about 60% of the spin density being located on the sulfur atom and only 35% on the copper atom.[17,18] This leads to the unusually small hyperfine coupling constant $A_{||}$ of copper of 161 MHz.[19] The same properties of the Cu^{2+} ion were also observed in several other blue copper proteins, which led to naming this type of Cu^{2+} ions as the *type 1* centers.

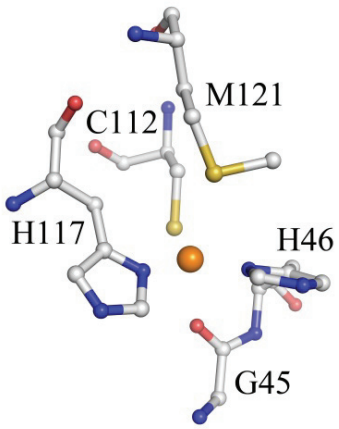


Figure 1.2.1. Coordination of the Cu^{2+} ion in azurin. The atoms of carbon, nitrogen, oxygen, and copper are shown as gray, blue, red, and orange spheres, respectively.

1.2.2 The Fe^{3+} ion in cytochrome P450cam

Cytochrome P450cam is an enzyme that catalyzes the hydroxylation of camphor at its 5-exo position (Figure 1.2.2a). It contains a heme group with an iron ion, whose valence state is 3+ in the resting state and varies between 2+ and 4+ during the catalytic reaction. In the resting state, a water molecule is bound to the Fe^{3+} ion as a sixth ligand resulting in an octahedral coordination sphere (Figure 1.2.2b) and a low-spin state ($S = 1/2$). Upon camphor binding, a water ligand is displaced yielding the penta-coordinated Fe^{3+} ion (Figure 1.2.2c). This change in coordination geometry results in the rearrangement of the energy levels of the iron orbitals and, in a spin state change from low-spin ($S = 1/2$) to high-spin ($S = 5/2$). This transition can be readily followed by continuous wave EPR (CW-EPR), which reveals the shift of the absorption signal from g = 2.45, 2.26, and 1.91 for the low-spin state to g = 7.85, 3.97, and 1.78 for the high-spin state.[20]

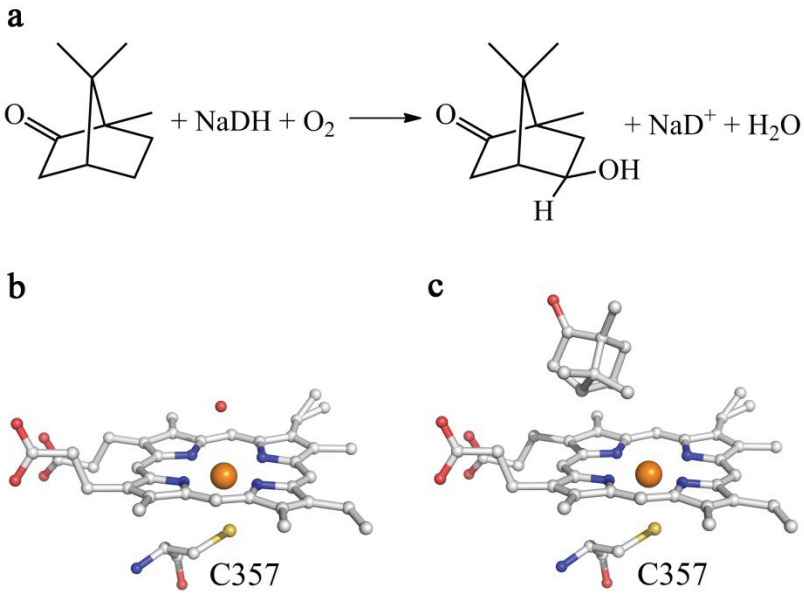


Figure 1.2.2. The catalytic reaction and the metal ion site of cytochrome P450cam. **a)** Schematic representation of the camphor hydroxylation reaction in cytochrome P450cam. **b)** Coordination of the Fe^{3+} ion in cytochrome P450cam in the absence of camphor. **c)** Coordination of the Fe^{3+} ion in cytochrome P450cam in the presence of camphor. The atoms of carbon, nitrogen, oxygen, and iron are shown as gray, blue, red, and orange spheres, correspondingly.

1.3 Site-directed spin labeling of biomolecules

Since the pioneering study of Hubbell and co-workers,[21] site-directed spin labeling (SDSL) has become a standard method to introduce paramagnetic probes into biomolecules, which allows investigation of the structure and dynamics of biomolecules by means of EPR. EPR measurements on spin-labeled biomolecules are often used to obtain information about

solvent accessibility of particular sites of biomolecules,[22] intra- or intermolecular distances between two spin-labeled sites[23,24] or a single spin-labeled site and an intrinsic paramagnetic center of a biomolecule.[14] SDSL is exceptionally useful for EPR-based distance measurements which are usually performed via Pulsed Electron-Electron Resonance (PELDOR)[25]. Using SDSL and PELDOR one can measure inter-spin distances up to 15 nm,[26] which matches nicely with the distance range of interest in biomolecules or biomolecular assemblies.

The most commonly used method for spin labeling of proteins utilizes the reactivity of the sulfhydryl group of cysteine residues. This approach usually requires that the target proteins possess cysteine residues only at the desired sites, which is achieved by means of site-directed mutagenesis. Usually, the sulfhydryl group is conjugated with the (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanesulfonothioate spin label (MTSSL)[27] by the formation of a disulfide bond (Figure 1.3.1a). The resulting spin-labeled side chain being commonly abbreviated as an R1 side chain (Figure 1.3.1b). The link between the piperidine-oxyl moiety and the protein backbone renders the R1 side chain flexible, thereby minimizing disturbances of the native fold of the protein. However, the large conformational space accessible for the R1 side chain poses often a challenge in the interpretation of PELDOR-measured distances: The PELDOR experiment provides the distance between the spin centres, which are localised on the NO groups of the R1 side chains, while it is the distance between the C_α atoms, which is of actual interest to the structural biologist. Due to the intrinsic flexibility and ~ 0.7 nm length of the R1 side chain, these two distances are often not easy to correlate.[28]

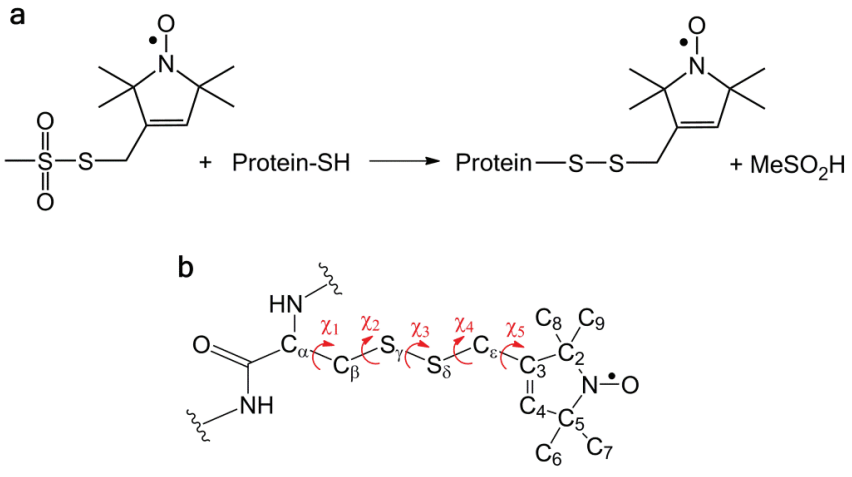


Figure 1.3.1. Spin labeling by MTSSL. **a)** The reaction of a cysteine with MTSSL. **b)** The structure of the R1 side chain. The spin is localized to more than 95% between the N and O atoms. The five dihedral angles χ_1-χ_5 corresponding to the rotatable bonds of the R1 side chain are shown by red arrows.

Besides MTSSL, a variety of other nitroxide spin labels with specific properties, such as stability under in-cell conditions or increased relaxation times, have been designed.[29] The use of unnatural amino acids has also been reported as a method for SDSL.[30,31]

As an alternative to nitroxides, spin labels based on Gd^{3+} chelates[32] and trityl radicals[33–36] are intensively investigated. Typical examples of these labels are MTS-ADO3A[37,38] (Figure 1.3.2a) and TAM[39] (Figure 1.3.2b). Compared to nitroxides, the Gd^{3+}-based labels offer an increased absolute sensitivity of PELDOR measurements at high frequencies/high magnetic fields and display stability towards the reductive environment found in cells.[40] Due to an extremely narrow EPR spectrum, trityl spin labels can also be advantageous with respect to an absolute sensitivity for the EPR-based distance measurements.[34] Moreover, trityl spins feature comparatively long relaxation times even at ambient temperatures, which allows using trityl spin labels for room-temperature EPR-based distance measurements.[35] Importantly, the development of alternative spin labels has also opened up a new field of PELDOR applications aiming at distance measurements between different types of labels; this is often referred to as "orthogonal" labeling.[41] The advantages of the "orthogonal" labeling are an improved sensitivity of the PELDOR measurements and the possibility to measure several distance constraints for one sample selectively.

Like in the case of proteins, a large variety of different methods for SDSL of nucleic acids has been proposed up to date.[42,43] These methods represent either spin labeling during oligonucleotide synthesis and post-synthetic spin labeling of pre-functionalized sites of the oligonucleotide. The first approach is limited to short sequences of 50-70 nucleotide residues that are obtained by the solid-phase synthesis. One of the typical spin labels here is **Ç** label[44] (Figure 1.3.2c). The second approach, in contrast, is not limited by the size of a nucleic acid and avoids subjecting the nitroxide to potentially reducing conditions during the oligonucleotide synthesis. A number of chemically modified nucleotide analogs have been proposed in the literature, revealing that the modification can be applied to all three structural units of oligonucleotides, the sugar moiety,[45,46] the phosphate backbone,[47,48] and the nucleobases.[49,50] A typical example of a post-synthetically attached spin label is the 3-(2-iodoacetamido)-proxyl spin label coupled to 4-thio uracil bases of DNA (Figure 1.3.2d).[51]

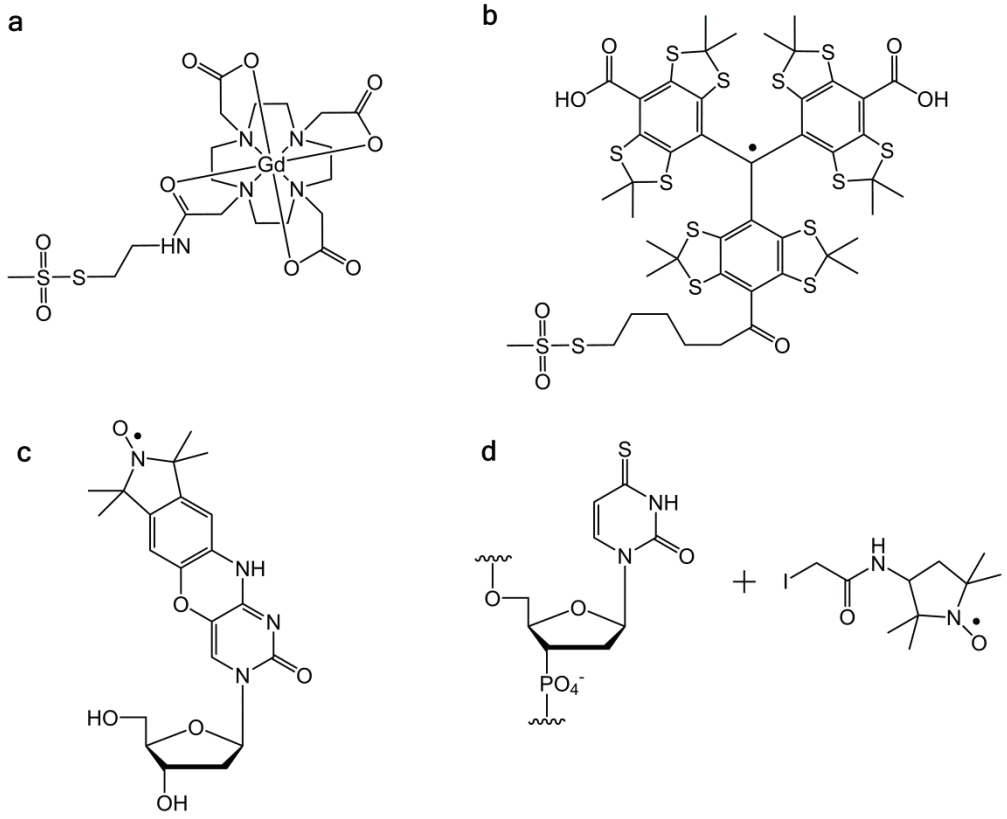


Figure 1.3.2. Examples of spin labels used for SDSL of biomolecules. **a)** MTS-ADO3A, **b)** TAM, **c)** Ç, and **d)** the proxyl label together with 4-thio uracil base of DNA.

1.4 Distance measurements via EPR

Electron Paramagnetic Resonance (EPR) is a technique based on the resonant absorption of microwave (m.w.) radiation by paramagnetic ions or molecules in a static magnetic field, a phenomenon discovered by Zavoisky in 1944.[52] The first EPR experiments were performed with continuous m.w. radiation at a constant frequency and a linearly changing magnetic field, named consequently as continuous wave EPR (CW-EPR) experiments. For many decades, CW-EPR has been successfully used to characterize paramagnetic species in crystals, chemical complexes, and biomolecules.[53] A major breakthrough in the field was the construction of pulsed EPR spectrometers, in which continuous radiation was replaced by m.w. pulses of microsecond to nanosecond duration. Although the first pulse EPR experiment was reported by Blume already in 1958,[54] the massive development and applications of pulsed EPR methods began after the release of the first commercial pulse EPR spectrometers in the 1980s . This yielded a number of pulsed methods, such as Electron Spin Echo Envelope Modulation (ESEEM), Hyperfine Sublevel Correlation (HYSCORE) and Electron-Nuclear Double Resonance (ENDOR), which enabled the investigation of relatively weak electromagnetic interactions between paramagnetic centers and their surroundings.[55] In the last 15 years, the development of EPR techniques to measure nanometer distances between

pairs of electron spins via dipolar couplings significantly increased the number of EPR applications to biological systems.[24] Two such techniques, known as Pulsed Electron-Electron Resonance (PELDOR, also known as DEER)[25] and Relaxation Induced Dipolar Modulation Enhancement (RIDME)[56], are also the subject of the present work.

Section 1.4 is organized as follows. First, a brief summary of the interactions underlying EPR-spectra is given followed by a short introduction of different EPR methods used to measure nanometer distances. A more detailed review of the basic theory and applications of the PELDOR and RIDME experiments concludes this section.

1.4.1 Spin Hamiltonian

Since electron and nuclear spins are quantum objects, a rigorous theory of EPR is based on quantum mechanics. In this theory, the energy of a spin system is described by a Hamiltonian operator (Hamiltonian). Using the Hamiltonian of a whole paramagnetic system for interpretation of experimental EPR data is usually an extremely complicated task. Instead, Abragam and Pryce have proposed an elegant way to simplify the general Hamiltonian by separating the energetic contributions involving the spin from all other contributions.[53] The obtained form of the Hamiltonian, called a static spin Hamiltonian, contains only phenomenological constants and spin coordinates described by the electron spin operator $\boldsymbol{S}$ and the m nuclear spin operators $\boldsymbol{I_k}$:

$$\begin{aligned} H_0 &= H_{EZ} + H_{ZFS} + H_{HF} + H_{NZ} + H_{NQ} = \\ &= \beta_e \boldsymbol{B_0}^T \boldsymbol{g} \boldsymbol{S} / \hbar + \boldsymbol{S}^T \boldsymbol{D} \boldsymbol{S} + \sum_{k=1}^{m} \boldsymbol{S}^T \boldsymbol{A_k} \boldsymbol{I_k} - \beta_n \sum_{k=1}^{m} g_{n,k} \boldsymbol{B_0}^T \boldsymbol{I_k} / \hbar + \sum_{I_k > 1/2} \boldsymbol{I_k}^T \boldsymbol{P_k} \boldsymbol{I_k}. \end{aligned} \tag{1.4.1}$$

The first term, H_{EZ}, describes the electron Zeeman interaction of the electron spin with an external magnetic field $\boldsymbol{B_0}$. The zero-field splitting of the energy levels for $S > 1/2$ is given by the second term, H_{ZFS}. The third term, H_{HF}, describes the hyperfine interaction of the electron spin with m nuclear spins. The fourth and fifth terms, H_{NZ} and H_{NQ}, sum up the energies of the nuclear Zeeman interactions and nuclear quadrupole interactions of m nuclei. The corresponding phenomenological constants are the g tensor of the electron spin $\boldsymbol{g}$, the zero-field splitting tensor $\boldsymbol{D}$, the tensors of the hyperfine interaction $\boldsymbol{A_k}$, the nuclear g factors $g_{n,k}$, and quadrupole interaction tensors $\boldsymbol{P_k}$. Note that, since the spatial degrees of freedom of the wavefunctions are assumed to be constants of motion, most of the phenomenological constants are second-rank tensors. A detailed description of the physical background of these constants can be found in the standard textbooks.[53,55] The constants β_e, β_n, and $\hbar$ are the Bohr magneton, the nuclear magneton, and the reduced Plank constant, correspondingly. The

symbol T indicates the transpose. All energies in Equation 1.4.1 are given in angular frequency units. Dependent on the number of interacting electron and nuclear spins, the state functions corresponding to the spin Hamiltonian H_0 span a Hilbert space with the dimension:

$$n_H = \prod_{k=1}^{N} (2J_k + 1), \tag{1.4.2}$$

where J_k corresponds to either the electron spin or one of the nuclear spins.

The main part of this work is concerned with the interactions between two electron spin centers, e.g. a transition metal ion and a nitroxide. These two centers are spatially well separated so that the magnetic interaction between them is much weaker than the Zeeman interaction of each of the spins with the static magnetic field. Such weekly interacting spins can be conveniently characterized by their individual spins S_A and S_B, and the energy of their interaction can be described by a sum of two terms, the exchange interaction term H_{ex} and the dipole-dipole interaction term H_{dd}. The complete spin Hamiltonian for such a two-spin system is then given by

$$H_0(S_A, S_B) = H_0(S_A) + H_0(S_B) + H_{exch} + H_{dd}, \tag{1.4.3}$$

where $H_0(S_A)$ and $H_0(S_B)$ denote the spin Hamiltonians for individual spins according to Equation 1.4.1. The exchange interaction term,

$$H_{exch} = \boldsymbol{S_A}^T \boldsymbol{J} \boldsymbol{S_B}, \tag{1.4.4}$$

is characterized by an exchange interaction tensor $\boldsymbol{J}$, which in general consists of isotropic and anisotropic parts. Except several specific cases for transition metal ions,[57] the anisotropic part of $\boldsymbol{J}$ can be neglected. The isotropic part is usually also negligibly small and becomes relevant only when two unpaired electrons are closer than 1.5 nm or are strongly delocalized.

The contribution of the dipole-dipole interaction into the spin Hamiltonian is given by

$$H_{dd} = \boldsymbol{S_A}^T \boldsymbol{D_{dd}} \boldsymbol{S_B} = \frac{\mu_0 \beta_e^2}{4\pi\hbar} \frac{g_A g_B}{r_{AB}^3} \left[\boldsymbol{S_A}^T \boldsymbol{S_B} - 3 \frac{(\boldsymbol{S_A}^T \boldsymbol{r}_{AB})(\boldsymbol{S_B}^T \boldsymbol{r}_{AB})}{r_{AB}^2} \right], \tag{1.4.5}$$

where $\boldsymbol{D_{dd}}$ is the dipole-dipole coupling tensor, μ_0 is the permeability of vacuum, $\boldsymbol{r}_{AB}$ is the vector connecting two electron spins, $r_{AB} = |\boldsymbol{r}_{AB}|$, and g_A and g_B are the g factors of the two electron spins. If the high-field approximation applies, meaning that both electron spins are quantized along the static magnetic field ($H_{EZ} \gg H_{ZFS}$), the scalar products in Equation 1.4.5 can be simplified using the orientation of $\boldsymbol{r}_{AB}$ with respect to the external magnetic field $\boldsymbol{B_0}$. Describing this orientation by means of two spherical angles θ and φ, one obtains the following expression for H_{dd}:

$$
\begin{aligned}
H_{dd} &= \frac{\mu_0 \beta_e^2}{4\pi\hbar} \frac{g_A(\theta,\varphi) g_B(\theta,\varphi)}{r_{AB}^3} [A + B + C + D + E + F] \\
A &= \boldsymbol{S}_{Az} \boldsymbol{S}_{Bz} (1 - 3\cos^2\theta) \\
B &= -\frac{1}{4} (\boldsymbol{S}_{A+} \boldsymbol{S}_{B-} + \boldsymbol{S}_{A-} \boldsymbol{S}_{B+}) (1 - 3\cos^2\theta) \\
C &= -\frac{3}{2} (\boldsymbol{S}_{A+} \boldsymbol{S}_{Bz} + \boldsymbol{S}_{Az} \boldsymbol{S}_{B+}) \sin\theta\cos\theta\, e^{-i\varphi} \\
D &= -\frac{3}{2} (\boldsymbol{S}_{A-} \boldsymbol{S}_{Bz} + \boldsymbol{S}_{Az} \boldsymbol{S}_{B-}) \sin\theta\cos\theta\, e^{i\varphi} \\
E &= -\frac{3}{4} \boldsymbol{S}_{A+} \boldsymbol{S}_{B+} \sin^2\theta\, e^{-2i\varphi} \\
F &= -\frac{3}{4} \boldsymbol{S}_{A-} \boldsymbol{S}_{B-} \sin^2\theta\, e^{2i\varphi},
\end{aligned}
\tag{1.4.6}
$$

where $\boldsymbol{S}_{Az}$ and $\boldsymbol{S}_{Bz}$ are the z-magnetization operators, $\boldsymbol{S}_{A+}$ and $\boldsymbol{S}_{B+}$ are the raising operators, and $\boldsymbol{S}_{A-}$ and $\boldsymbol{S}_{B-}$ are the lowering operators of the two spins. If the dipolar coupling constant,

$$
\omega_{dd} = \frac{\mu_0 \beta_e^2}{4\pi\hbar} \frac{g_A(\theta,\varphi) g_B(\theta,\varphi)}{r_{AB}^3}, \tag{1.4.7}
$$

is much smaller as than the resonance frequencies of the individual spins, ω_A and ω_B (the eigenvalues of $H_0(S_A)$ and $H_0(S_B)$), $\omega_{dd} \ll \omega_A$ and $\omega_{dd} \ll \omega_B$, the non-secular terms C-F in Equation 1.4.6 can be neglected. Furthermore, if the difference between the resonance frequencies of the individual spins is greater than dipolar coupling constant, $\omega_{dd} \ll |\omega_A - \omega_B|$, the pseudo-secular term B can also be neglected. Applying these considerations and neglecting the anisotropy of the $\boldsymbol{g}$ tensors, H_{dd} can be written as

$$
H_{dd} = \omega_{dd} (1 - 3\cos^2\theta) \boldsymbol{S}_{Az} \boldsymbol{S}_{Bz}. \tag{1.4.8}
$$

Since most of the biomolecular probes are solutions, it is important to note that the rotation of spin-carrying molecules with the frequency exceeding the dipolar coupling constant averages the angular term in Equation 1.4.8 to zero. For this reason, systems under study have to be immobilized. In the case of biomolecules, this is normally achieved by freezing the sample into a homogeneous glass. Other methods to prevent fast tumbling include attachment of the biomolecule to a solid support or using viscous solvents.[35,36]

In a typical case of a frozen sample with random orientations of $\boldsymbol{r}_{AB}$ vectors with respect to the static magnetic field $\boldsymbol{B}_0$, the angular dependence of H_{dd} gives rise to a distinct dipolar spectrum known as a Pake doublet. A typical shape of the Pake doublet is shown in Figure 1.4.1. It contains two singularities, corresponding to $\theta = 0°$ and $\theta = 90°$ and appearing at the frequencies of ω_{dd} and $2\omega_{dd}$, respectively. These frequencies are often called in the literature as parallel ($\omega_{\parallel}$) and perpendicular ($\omega_{\perp}$) components of the Pake doublet.

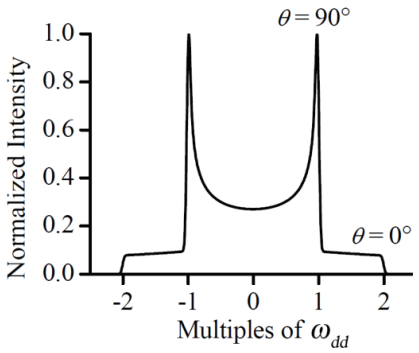


Figure 1.4.1. The Pake doublet. The two singularities correspond to $\theta = 0°$ and $\theta = 90°$.

Different EPR methods have been developed to measure dipolar coupling constants, depending on the spectral and dynamical properties of the paramagnetic centers, as well as the distance between them. In the following sections, these methods will be briefly described, mentioning their application range, advantages, and limitations.

1.4.2 CW-EPR

A general description of the continuous wave EPR (CW-EPR) experiment can be found in a number of excellent textbooks.[58,59] Here, the basic principles of the CW-EPR experiment will be discussed on the example of the Cu^{2+} ion of azurin (Figure 1.2.1). The Cu^{2+} ion represents a spin system, which consists of a single unpaired electron ($S = 1/2$) and a magnetic Cu nucleus ($I = 3/2$). In the CW-EPR experiment, the Cu^{2+} ion is subjected to a static magnetic field $\boldsymbol{B_0}$ so that its energy is described by the spin Hamiltonian given by Equation 1.4.1. The electron Zeeman and hyperfine interaction terms of this Hamiltonian are much larger than the other terms for the Cu^{2+} ion, which makes it sufficient to use the shortened version of the spin Hamiltonian,

$$H_0 = \beta_e \boldsymbol{B_0}^T \boldsymbol{gS}/\hbar + \boldsymbol{S}^T \boldsymbol{AI}. \tag{1.4.9}$$

Note that, for simplicity, only the hyperfine coupling of the electron spin to the Cu nuclear spin is considered, whereas the hyperfine coupling to the nuclear spins of the nitrogen ligands is neglected. The $\boldsymbol{g}$ and $\boldsymbol{A}$ tensors of the Cu^{2+} ion are anisotropic. In the molecular coordinate system of the Cu^{2+} ion (Figure 1.4.2a), the anisotropy can be described by three values for each tensor: $g_{xx}, g_{yy}, g_{zz} = 2.032, 2.047, 2.253$ and $A_{xx}, A_{yy}, A_{zz} = 30, 18, 172$ MHz.[60,61]

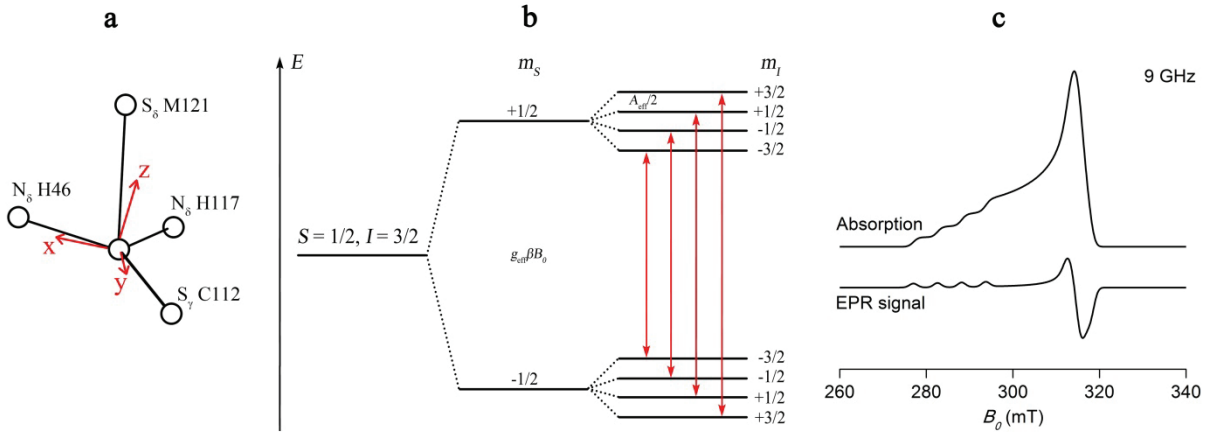


Figure 1.4.2. CW-EPR on the Cu^{2+} ion of azurin. **a)** The molecular coordinate system of the Cu^{2+} ion is shown with respect to the ligands. Adapted from Ref. [60]. **b)** The energy diagram of the Cu^{2+} ion and the corresponding absorption spectrum. The allowed EPR transitions are shown by red arrows. c) The powder-averaged spectrum of the Cu^{2+} ion and its first derivative, representing the CW-EPR signal. The spectra were simulated by means of the program EasySpin[62] using $\omega_{mw}/2\pi = 9$ GHz (X-band), $g = [2.032, 2.047, 2.253]$, $A = [30, 18, 172]$ MHz and an inhomogeneous linewidth of 50 MHz.

Solving the Schrödinger equation for the spin Hamiltonian given by Equation 1.4.9 yields eight eigenstates $|m_S, m_I\rangle$ with energies

$$E(m_S, m_I) = g_{eff}\beta_e B_0 m_S / \hbar + A_{eff} m_S m_I, \tag{1.4.10}$$

where $m_S = \pm 1/2$, $m_I = \pm 1/2$ or $\pm 3/2$, g_{eff} and A_{eff} are the effective g factor and hyperfine coupling constant for the particular orientation of Cu^{2+} molecular axes with respect to the magnetic field $\boldsymbol{B_0}$. These energy levels are plotted in dependence of the magnetic field B_0 in Figure 1.4.2b. The transitions between these energy levels with the selection rules $\Delta m_S = \pm 1$ and $\Delta m_I = 0$ can be induced by applying m.w. radiation, which creates a circularly polarized magnetic field $\boldsymbol{B_1}$ perpendicular to the static magnetic field $\boldsymbol{B_0}$. The frequency of the m.w. radiation ω_{mw} needs to satisfy the resonance condition,

$$\omega_{mw} = g_{eff}\beta_e B_0 / \hbar + A_{eff} m_I. \tag{1.4.11}$$

This equation yields four possible transitions for the Cu^{2+} ion. They are depicted as arrows in Figure 1.4.2b. In the CW-EPR experiment, ω_{mw} is kept constant and the magnetic field B_0 is linearly incremented. Thus, the transitions are observed as four absorption peaks appearing at four different magnetic fields.

Up to now all calculations were done for a single Cu^{2+} ion, whereas the actual CW-EPR spectrum is usually acquired on 10^{15}-10^{21} such ions. If the corresponding azurin molecules are diluted in frozen solution, the Cu^{2+} spin center has random orientations with respect to the static magnetic field $\boldsymbol{B_0}$. In this case, the EPR spectrum represents a superposition of Cu^{2+} spectra with all possible values of g_{eff} and A_{eff}. This spectrum is called a powder-averaged spectrum (Figure 1.4.2c). In the CW-EPR experiment, this spectrum is recorded as the first

derivative of the absorption spectrum (Figure 1.4.2c). The advantages of such acquisition scheme are a reduction of noise that does not oscillate with the modulation frequency and a better resolution of the non-resonant m.w. absorption that does not depend on the magnetic field. Furthermore, the derivative spectrum is better resolved than the absorption spectrum.
The spectral lines observed in the CW-EPR spectrum of Cu^{2+} are homogeneously and inhomogeneously broadened. The source of the homogeneous broadening is fluctuating fields originating e.g. from electron and nuclear spin flips. The inhomogeneous broadening stems from the anisotropy of the $\boldsymbol{g}$ and $\boldsymbol{A}$ tensors, unresolved hyperfine structures, and inhomogeneity of the external static magnetic field $\boldsymbol{B_0}$.

1.4.3 Distance measurements via CW-EPR

Since energy levels of electron spins are in general determined by all interactions included in the spin Hamiltonian, a CW-EPR spectrum contains information about all phenomenological constants of Equations 1.4.1 and 1.4.3. However, due to the anisotropy of dominant interactions and inhomogeneity of the EPR probes at the macroscopic level, smaller interactions are often not resolved and contribute only to the width of the EPR spectrum. Thus, the determination of dipolar coupling constants by CW-EPR requires that the dipolar coupling constant is larger than the intrinsic EPR linewidth of at least one of the paramagnetic centers involved. Usually, the dipolar splitting constant is smaller than the intrinsic EPR linewidth and leads only to a line broadening. In these cases, deconvolution methods[63] or a multiple-parameter fitting[64,65] can be used to disentangle the dipolar contribution to the linewidth from other contributions. However, these methods strongly depend on the quality of the EPR data and are limited for nitroxides to inter-spin distances in the range of 1-2 nm.[66] To reliably measure the inter-spin distances above 2 nm, a relatively weak dipole-dipole interaction has to be separated from the other interactions influencing the EPR spectrum. This can be achieved using various pulsed EPR techniques.

1.4.4 Pulsed EPR

An in-depth description of pulsed EPR is provided in the textbook of Schweiger and Jeschke.[55] The authors of this book mention two different ways to interpret pulsed EPR experiments. One of them, known as a classical description, is based on a classical motion of a magnetization of a spin ensemble. A more rigorous description is provided by the quantum-mechanical formalism, which deals with a time evolution of a spin density matrix. Here, the

first way will be briefly outlined, because it provides a more intuitive picture of pulsed EPR and is sufficient for the purposes of the present work.

The central quantity in the classical description is the total magnetic moment of an ensemble of N electron spins known as the magnetization $\boldsymbol{M}$,

$$\boldsymbol{M} = \frac{1}{V}\sum_{i=1}^{N} \boldsymbol{\mu}_i, \tag{1.4.12}$$

where V is the volume occupied by the spin ensemble, and $\boldsymbol{\mu}_i$ are the magnetic moments of individual spins

$$\boldsymbol{\mu}_i = -g\beta_e \boldsymbol{S}_i. \tag{1.4.13}$$

In a pulsed EPR experiment, the motion of the magnetization $\boldsymbol{M}$ is governed by two magnetic fields, the static field $\boldsymbol{B_0}$ and perpendicular to it the circularly polarized field $\boldsymbol{B_1}$. In contrast to CW-EPR, the $\boldsymbol{B_1}$ field is created not by continuous m.w. radiation, but by high-power m.w. pulses of nanosecond length. In thermal equilibrium, $\boldsymbol{M}$ is aligned parallel to $\boldsymbol{B_0}$. In the case $\boldsymbol{M}$ is shifted from its equilibrium direction, $\boldsymbol{M}$ precesses about $\boldsymbol{B_0}$ with the Larmor frequency

$$\omega_0 = \frac{g\beta_e}{\hbar} B_0. \tag{1.4.14}$$

When a m.w. pulse is applied and its frequency ω_{mw} is set in resonance with the Larmor frequency ω_0, $\omega_{mw} = \omega_0$, $\boldsymbol{M}$ experiences a precession about $\boldsymbol{B_1}$ with the frequency

$$\omega_1 = \frac{g\beta_e}{\hbar} B_1. \tag{1.4.15}$$

The effect of a m.w. pulse on the dynamics of $\boldsymbol{M}$ is best shown using a rotating frame, in which $\boldsymbol{B_0}$ is aligned along the z-axis and $\boldsymbol{B_1}$ is directed along the x-axis (Figure 1.4.3a). The spin system can be assumed to be initially in equilibrium so that $\boldsymbol{M}$ is parallel to the z-axis, $\boldsymbol{M} = [0, 0, M_0]$. Applying a m.w. pulse of length t_p induces the precession of $\boldsymbol{M}$ about the x-axis. As a result, the projection of $\boldsymbol{M}$ on the frame axes at time t_p is given by

$$\begin{aligned} M_x(t_p) &= 0, \\ M_y(t_p) &= -M_0 \sin\beta, \\ M_z(t_p) &= M_0 \cos\beta, \end{aligned} \tag{1.4.16}$$

where $\beta = \omega_1 t_p$ is the flip angle of $\boldsymbol{M}$ about the x-axis. When β equals $\pi/2$, the corresponding m.w. pulse is called $\pi/2$-pulse. Here, the $\pi/2$-pulse rotates $\boldsymbol{M}$ from the z direction into $-y$ direction. A similar rotation of $\boldsymbol{M}$ about y, $-x$, and $-y$ directions can be achieved by shifting the phase of the m.w. pulse by $\pi/2$, π, and $3\pi/2$, correspondingly. Thus, any desired rotation of $\boldsymbol{M}$ can be realized with appropriately timed and phased m.w. pulses.

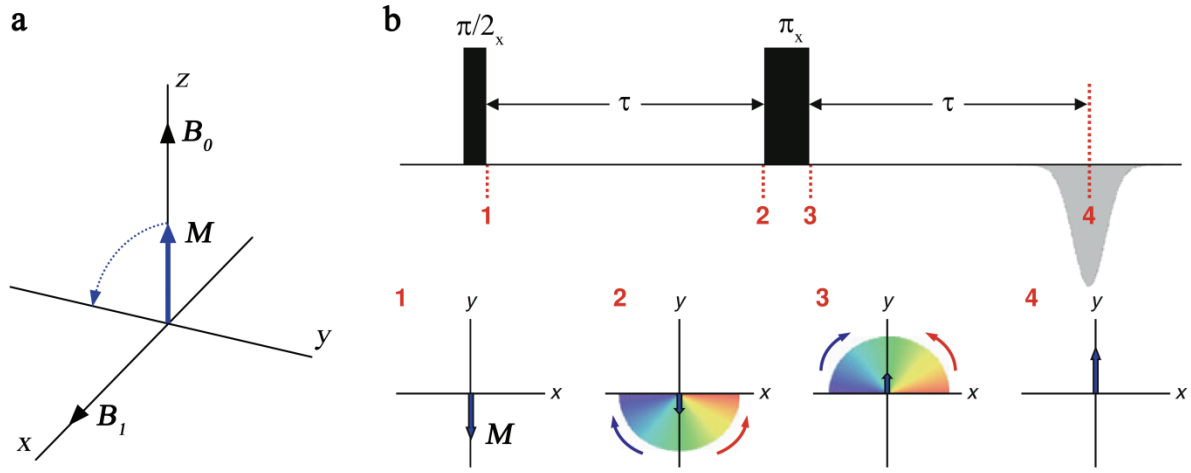


Figure 1.4.3. The basics of pulsed EPR. **a)** The rotating frame used to describe the motion of the magnetization ***M*** in the presence of the static magnetic field $\boldsymbol{B_0}$ and the circularly polarized magnetic field $\boldsymbol{B_1}$. **b)** Hahn echo experiment. Top: The pulse sequence. Bottom: The orientation of ***M*** in the *xy* plane of the rotating frame at different stages of the Hahn echo experiment. Adapted from Ref. [67].

If the resonance condition $\omega_{mw} = \omega_0$ is not fulfilled for all spins taking part in the pulsed EPR experiment, an additional precession of ***M*** about the *z*-axis with the frequency $\Omega_0 = \omega_0 - \omega_{mw}$ occurs after the m.w. pulse. Moreover, a free evolution of the spin system after the m.w. pulse is accompanied by the return of ***M*** to its equilibrium state through so-called relaxation processes. The relaxation processes include an exponential decay of M_x and M_y components with the characteristic time T_2, called transverse relaxation time, as well as an exponential increase of the M_z component with the characteristic time T_1, called longitudinal relaxation time. Taking all this into account, the components of ***M*** at time *t* after the m.w. pulse are

$$\begin{aligned} M_x(t_p+t) &= M_0 \sin\beta \sin(\Omega_0 t)\exp(-t/T_2), \\ M_y(t_p+t) &= -M_0 \sin\beta \cos(\Omega_0 t)\exp(-t/T_2), \\ M_z(t_p+t) &= M_0\left[1-(1-\cos\beta)\exp(-t/T_1)\right]. \end{aligned} \qquad (1.4.17)$$

The M_x and M_y components of ***M*** are coupled to the m.w. resonator and, therefore, can be measured in the pulsed EPR experiment. Both components are usually measured simultaneously by means of a quadrature detection scheme, yielding an EPR signal that is proportional to $M_y - iM_x$. Thus, the EPR signal detected after a single m.w. pulse is determined by

$$V(t) \propto \exp(i\Omega_0 t)\exp(-t/T_2). \qquad (1.4.18)$$

This signal is known as the free induction decay (FID). FIDs are commonly used in NMR spectroscopy. However, due to a relatively long dead-time of most EPR spectrometers (~100 ns at X-band), a spin echo is usually detected in pulsed EPR experiments instead of

FID. The simplest representative of a spin echo, the primary echo (or Hahn echo), is observed after the pulse sequence $\pi/2_x$ - τ - π_x - τ. The echo formation in this sequence is explained in Figure 1.4.3b. The first $\pi/2_x$-pulse rotates $\boldsymbol{M}$ in the -y direction. During the first free evolution time τ spin packets with different Larmor frequencies gain a phase shift due to their different angular precession. In the rotating frame, the spin packets fan out according to their resonance offset Ω_0 and, consequently, the magnitude of M_y is decreased. The π_x-pulse inverts the sign of the $\boldsymbol{M}$ precession in the xy-plane. After the second free evolution time τ the phases of the different spin packets are refocused and M_y is maximized up to the value that can be measured. Note that the increase of the interval τ leads to the decrease of the echo signal due to the T_2 relaxation process mentioned above.

In the next sections, more elaborate pulse sequences, which are however based on the same principles as discussed here, will be considered.

1.4.5 Distance measurements via pulsed EPR

Pulsed EPR experiments allow to refocus all static inhomogeneous broadening contributions and to recover the much narrower homogeneous linewidth, given by the transversal relaxation rate $1/T_2$, thus tremendously increasing the spectral resolution as compared to CW-EPR. Furthermore, the contribution of different electromagnetic interactions to the EPR signal can be separated by choosing an appropriate pulse sequence. In particular, the dipole-dipole and exchange interactions between two electron spins can be isolated from other interactions, enabling the inter-spin distance measurements in the range 1.5 – 15 nm. A number of pulsed EPR techniques capable of recovering inter-spin distances through dipolar coupling constants were developed. These include single frequency techniques, e.g. double-quantum coherence EPR (DQC),[68] single frequency technique for refocusing dipolar couplings (SIFTER),[69] '2+1',[70,71] relaxation induced dipolar modulation enhancement (RIDME),[56] T_1- and T_2-relaxation based methods,[72–74] as well as a double frequency technique called pulsed electron-electron double resonance (PELDOR)[25] or, alternatively, double electron-electron resonance (DEER)[75]. The single frequency techniques are in general technically less demanding, but they are more limited to spin systems with particular spectroscopic properties as compared to PELDOR. A detailed description of the advantages and limitations of these techniques can be found in several review articles.[24,76,77] Here, the important features of the most commonly used techniques will be briefly mentioned with the main emphasis on the applicability of each technique to paramagnetic metal ions.

The most commonly used method to measure dipolar coupling constants is PELDOR. The original 3-pulse PELDOR pulse sequence[25] is identical to Hahn's spin echo double resonance (SEDOR) experiment[78] to detect the coupling between nuclear spins. Nowadays, the 4-pulse PELDOR sequence[75] (Figure 1.4.4a) is used, because it is dead-time free as compared to the original sequence. Recently, the 5-pulse PELDOR sequence[79] and the 4-pulse PELDOR sequence with coherent,[80] composite,[81] and adiabatic[82] pulses were also established. However, only the standard four-pulse sequence was used in this thesis. This sequence consists of m.w. pulses at two different frequencies, which are set to resonance frequencies of both spin centers involved in the dipole-dipole interaction. The optimal performance of PELDOR requires that the m.w. pulses at different frequencies should have non-overlapping excitation windows. For the typical pulse lengths of 16-32 ns, the overlap of the excitation windows can be minimized when the resonance frequencies of the spin centers differ by more than 80 MHz. Along with that, the difference in the resonance frequencies of the spin centers should not exceed the bandwidth of the EPR resonator (~160 MHz for the fully overcoupled Bruker QT-II resonator) to avoid strong decrease of m.w. field strength for the m.w. pulses as well as reduction of signal-to-noise ratio for the PELDOR signal.

When the difference between the resonance frequencies is smaller than the bandwidths of the m.w. pulses applied, single frequency techniques such as DQC and SIFTER are more convenient than PELDOR. In DQC the dipolar coupling is separated from other contributions to the spin Hamiltonian by introducing a double quantum filter in analogy to liquid state NMR.[83] The most commonly used 6-pulse DQC sequence (Figure 1.4.4b) can be portioned into four periods: preparation, evolution, mixing, and detection. During the "preparation" period double-quantum coherence (DQC) is created by means of the 3-pulse sequence $\pi/2 - \tau_1 - \pi - \tau_1 - \pi/2$. This coherence is then refocused by the sequence $T - \pi - T$ of the "evolution" period, which serves together with the appropriate phase cycle as a double-quantum filter in the coherence pathway of the observable signal and suppresses all contributions not related to dipolar coupled spin pairs. During the "mixing" period DQC is converted by the fifth pulse into the single-quantum coherence (SQC). Finally, the SQC is refocused by the sixth pulse to form a detected echo. The echo signal is recorded as a function of $(\tau_2 - 2\tau_1)$, where τ_1 is incremented by steps Δt and τ_2 is decremented by the same steps. The optimal performance of the DQC experiment is achieved when the coupled spin centers are fully excited by the m.w. pulses. This requirement is however very challenging because most of the paramagnetic species have spectral widths which exceed the excitation bandwidths of the commonly used m.w. pulses of 8 – 32 ns length. Nevertheless, these experiments have

been successfully carried out in a number of paramagnetic systems including nitroxide and trityl radicals.[66]

Another single frequency technique, SIFTER, uses the same principle to refocus the dipolar coupling as the solid-echo sequence[84] known from NMR. The original solid-echo sequence cannot be used in EPR because it does not completely refocus inhomogeneity due to a distribution of *g* factors or hyperfine couplings. This problem is solved in SIFTER by introducing two more refocusing π pulses. The sequence $(\pi/2)_x - \tau_1 - (\pi)_x - \tau_1 - (\pi/2)_y - \tau_2 - (\pi)_x - \tau_2 - echo$ generates a solid echo for $\tau_1 = \tau_2$. The *g* and hyperfine inhomogenities are refocused by the two π pulses for any choice of τ_1 and τ_2, so that a variation of $(\tau_2 - \tau_1)$ results in echo modulation which is exclusively due to the dipolar coupling. This variation is done by incrementing delay τ_2 by steps Δt and decrementing delay τ_1 by the same steps. Note that, similar to the DQC experiment, the SIFTER experiment requires complete excitation of coupled spins. When this requirement cannot be fulfilled, the SIFTER signal can have contributions from unwanted coherence pathways. To avoid this difficulty, the use of broadband adiabatic m.w. pulses for SIFTER has been recently proposed.[85]

All pulsed EPR techniques described so far require that both coupled spins exhibit slow relaxation rates and can be sufficiently excited by the m.w. pulses. In contrast, relaxation based methods need only one of the spin centers to be relaxing slow enough to be detected, whereas the dipolar coupling is affecting the detected spin through the relaxation of the second spin. In principle, the mechanism altering the spin state of the coupled spin is not of importance – be it a m.w. pulse or longitudinal relaxation. The only difference is that the inversion of the spin by relaxation is not coherent, in contrast to the inversion by a m.w. pulse. If the relaxation processes of the two coupled spins are not correlated, the total relaxation rate of the detected spin will be a product of the intrinsic relaxation and the dipolar relaxation rate. Thus, the dipolar relaxation can be isolated, in case that the intrinsic relaxation of the detected spin can be studied independently. Depending on the value of the longitudinal relaxation time of the fast relaxing spin, the effect of dipolar coupling on the transversal or longitudinal relaxation rate ($1/T_2$ or $1/T_1$) of the detected spin can be detected. The relaxation rates $1/T_2$ and $1/T_1$ are measured using the 2-pulse ESEEM (Figure 1.4.4d) and the inversion recovery sequences (Figure 1.4.4e), respectively. The disadvantage of relaxation methods is that the dipolar effect leads only to an additional exponential decay of the signal and not to a coherent oscillation, because of the statistic nature of the spin-flip processes. Separation of this contribution from intrinsic relaxation mechanisms is usually not trivial.[24]

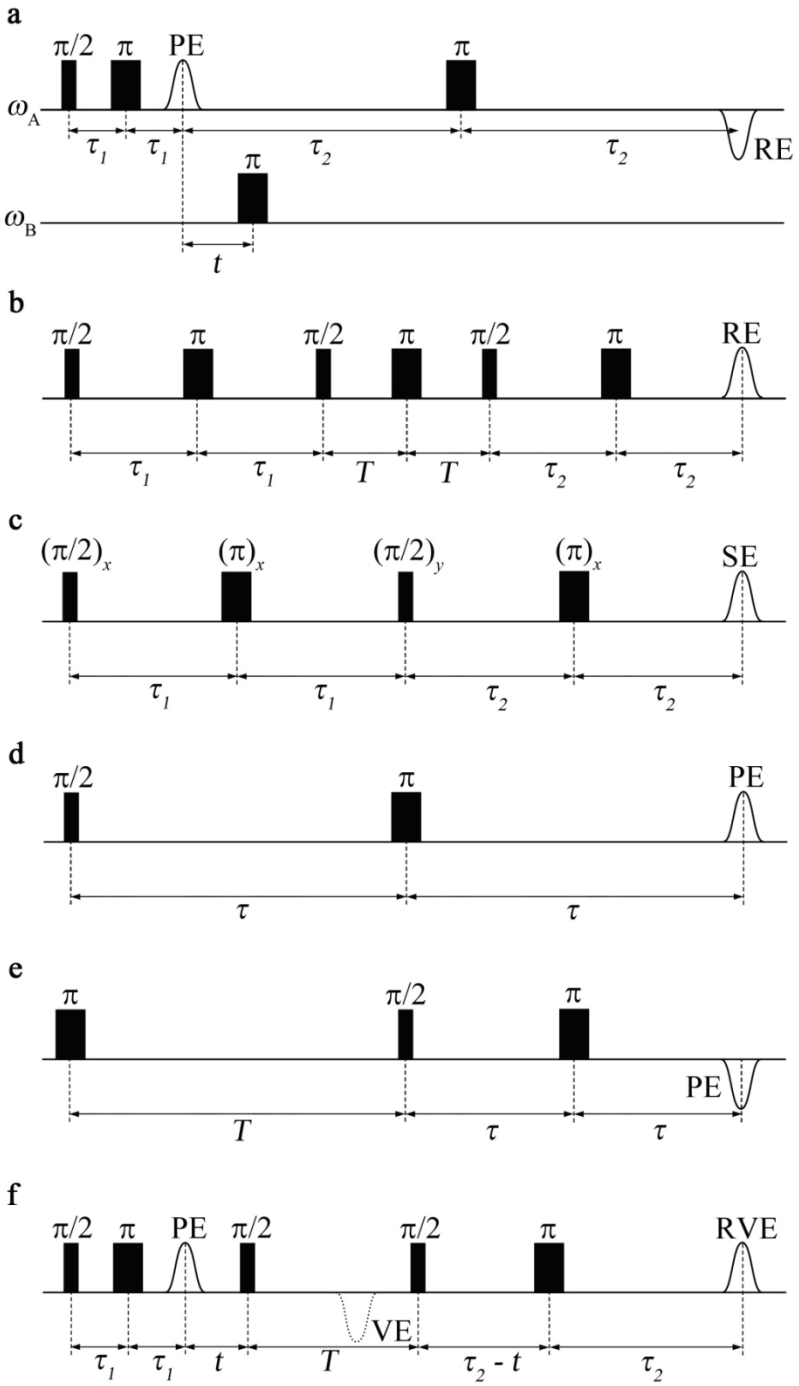


Figure 1.4.4. Pulsed EPR sequences for distance measurements. **a)** 4-pulse PELDOR, **b)** 6-pulse DQC, **c)** SIFTER, **d)** 2-pulse ESEEM, **e)** inversion recovery, and **f)** 5-pulse RIDME. The positions of the primary (PE), refocused (RE), solid (SE), virtual (VE), and refocused virtual (RVE) echoes are marked in the pulse sequences.

A further method for accessing dipolar couplings between a slow and a fast relaxing species is RIDME. In contrast to the relaxation methods described above, this method yields time traces in which the dipolar coupling manifests itself as an oscillation, similar to the PELDOR time traces. Since the pioneering work of Kulik and co-workers,[56] the RIDME experiment has been significantly advanced. The disadvantage of the original 3-pulse RIDME sequence is its dead time, which obscures the initial part of the signal and, therefore, complicates the

extraction of the distance distribution. The introduction of the 4-pulse sequence[86] allowed one to reduce the dead time, and the 5-pulse sequence[87] (Figure 1.4.4f) recently proposed by the lab of Huber has been shown to be dead time free. The 5-pulse sequence will be discussed in detail in Section 1.10.

To summarize, pulsed EPR offers a large arsenal of different techniques to measure the distances between spin centers. However, not all these techniques are well suited for the applications on metal ions. The main limitation stems from the fact that many paramagnetic metal ions have a large spectral width. This makes it impossible to fulfill the optimal conditions for the DQC and SIFTER experiments. Moreover, the PELDOR experiment becomes orientation selective, which poses significant challenges for the analysis of its experimental data. This problem is discussed in detail in the following sections. In contrast, the relaxation-based techniques are well suited for relatively fast relaxing metal ions, especially, when they are coupled to slower relaxing spin centers such as nitroxides. Among the relaxation-based techniques, RIDME seems to be the most attractive technique, because of its ability to provide the signals modulated by dipolar frequencies.

1.4.6 PELDOR

The scheme of the 4-pulse sequence of the PELDOR experiment is shown in Figure 1.4.4a. The pulses in this sequence are applied at two different m.w. frequencies called the detection frequency (ω_A) and the pump frequency (ω_B). The first two m.w. pulses with a fixed inter-pulse separation τ_1 create a primary spin echo (Hahn echo) from those spins that are in resonance with the detection frequency (A spins). The fourth π-pulse is applied to the same spins at a time τ_2 after the primary echo and creates a refocused spin echo at a time τ_2 after this pulse. The introduction of the third π-pulse at the pump frequency (pump pulse) flips the spins that are in resonance with the pump frequency (B spins). Due to the dipole-dipole interaction between the A and B spins, the flip of the B spins changes the local magnetic field at the A spins. As a consequence, the Larmor frequency of the A spins is shifted by

$$\omega_{AB} = \omega_{dd}\left(1 - 3\cos^2\theta\right), \tag{1.4.19}$$

where ω_{dd} and θ are defined by Equations 1.4.6-1.4.8. Depending on the sign of the frequency shift, the A spins accumulate a phase shift of $\pm\omega_{AB}t$, where t defines the time position of the pump pulse with respect to the primary echo. The total magnetization of these two A spin sub-ensembles oscillates then as

$$\exp(i\omega_{AB}t) + \exp(-i\omega_{AB}t) \propto \cos(\omega_{AB}t). \tag{1.4.20}$$

Thus, for different t values the primary echo is refocused with time-dependent phases, which give rise to the modulation of the intensity of refocused echo with the dipolar frequencies. This dependence is recorded in the PELDOR experiment. The corresponding PELDOR signal is called a PELDOR time trace.

The time dependence of the PELDOR time trace can be expressed by a product of two contributions,

$$V_{PELDOR}(t) = V_{intra}(t)V_{inter}(t). \tag{1.4.21}$$

$V_{intra}(t)$ denotes the contribution of the dipolar interaction between spin centers within the same molecule, whereas $V_{inter}(t)$ corresponds to the echo decay caused by the intermolecular distribution of the spin centers in the sample. In the case of a homogenous distribution of molecules in the sample, $V_{inter}(t)$ is a monoexponential decay and can be readily removed during the data analysis. $V_{intra}(t)$ describes the part of the PELDOR time traces which is modulated by dipolar frequencies. For a single spin pair, $V_{intra}(t)$ is given by

$$V_{intra}(t) = V_0\{1 - \lambda[1 - \cos(\omega_{AB}t)]\}. \tag{1.4.22}$$

Here, V_0 is the A spin echo intensity at $t = 0$, and λ is the probability of the inversion of the B spin by the pump pulse, also called as a modulation depth parameter. Note that V_0, λ, and ω_{AB} depend on the orientation of the magnetic field $\boldsymbol{B_0}$ with respect to the A spin, B spin, and distance vector $\boldsymbol{r_{AB}}$. For N conformers of the spin pair $V_{intra}(t)$ can be written as a sum,

$$V_{intra}(t) = \sum_{i=1}^{N} V_{0i}\{1 - \lambda_i[1 - \cos(\omega_{AB}^i t)]\}. \tag{1.4.23}$$

For disordered powder samples, such as frozen solutions, Equation 1.4.23 can be averaged over random orientations of the magnetic field,

$$V_{intra}(t) = \sum_{i=1}^{N} \int_0^{2\pi}\int_0^{\pi} V_{0i}(\varphi,\psi)\{1 - \lambda_i(\varphi,\psi)[1 - \cos(\omega_{AB}^i(\varphi,\psi)t)]\}\sin\varphi d\varphi d\psi, \tag{1.4.24}$$

where the polar angles φ and ψ describe the orientation of the magnetic field in the laboratory coordinate frame.

If the correlation between the orientations of the A spins, the B spins, and the distance vectors can be neglected, the integration over φ and ψ in Equation 1.4.24 can be substituted by the integration over all dipolar angles θ with the probability density function $sin(\theta)$, and the summation over N conformation can be replaced by the integration over the inter-spin distances with the probability density function $D(r)$,

$$V_{intra}(t) = \langle V_0 \rangle \left\{ 1 - \langle \lambda \rangle \int_0^{\infty} D(r) \int_0^{\pi/2} [1 - \cos(\omega_{AB}(r,\theta)t)] \cdot \sin\theta \, d\theta dr \right\}, \quad (1.4.25)$$

Here, $\langle V_0 \rangle$ and $\langle \lambda \rangle$ are the total A spin echo intensity at $t = 0$ and the modulation depth parameter for the whole ensemble of spin pairs, respectively. Note that Equation 1.4.25 usually holds for spin labels with flexible molecular tethers. For this case, the distance distribution $D(r)$ can be deduced from $V_{intra}(t)$ using Tikhonov regularization.[88] This algorithm is implemented in the program DeerAnalysis,[89] which is at present a common tool for obtaining distance distribution from the raw PELDOR data described by Equation 1.4.25.

In the opposite case, when the correlations between the orientations of the A spins, the B spins and the distance vectors are significant, Equation 1.4.24 has to be explicitly used in the analysis. This case is referred in the literature as orientation selective PELDOR. It is usually realized for rigid spin centers such as rigid spin labels or metal centers tightly bound to biomolecules. The intricacy of the orientation selective PELDOR data analysis arises from the fact that the correlation between the orientations of the spins and the distance vectors are usually not known *a priori*. The existing approaches to predict this correlation and, consequently, to extract the distance distribution function $D(r)$ from the orientation selective PELDOR data are discussed in the next section.

1.4.7 Orientation-selective PELDOR

In recent years, several approaches were proposed for the analysis of orientation selective PELDOR data. In most of these works, PELDOR time traces were collected at several magnetic fields and/or frequency offsets (ω_A - ω_B) in order to probe different orientations of the spin centers. If the majority of the orientations are probed, summing up all time traces can reduce the effect of orientation selection. In this case, the easiest way to obtain the inter-spin distance is to analyze the resulting averaged time trace by Tikhonov regularization.[90] This approach usually provides a good estimate of the inter-spin distance. However, the width of the distance distribution cannot be determined with high accuracy and all information about the orientation of spin centers is lost.

In a more elaborate approach, a molecular model defining the relative orientation of the coupled spin centers is constructed and used to simulate PELDOR time traces. Several models were used to simulate PELDOR time traces of nitroxide biradicals,[90–95] amino acid biradicals,[96,97] trityl-nitroxide biradicals,[34] and metal-nitroxide as well as metal-metal complexes.[98–100] In most studies, the initial model of the spin pair was based on the data obtained from X-ray diffraction measurements and/or DFT and MD simulations. Then, the

geometric parameters of the model were varied until the simulated PELDOR time traces matched the experimental ones. A significant drawback of this method is that a large space of geometric parameters has to be scanned until the optimal geometry of the model is found. Unless an optimization algorithm is applied for this, this task becomes very time-consuming and it is not clear whether the solution obtained is unique. In the papers of Abé *et al.*[101] and Tkach *et al.*[102], the number of optimized geometric parameters could be reduced down to 3 and 6 due to the high rigidity of the spin system studied. This allowed exploration of the whole possible range of each optimization parameter within a reasonable time and, consequently, solving the problem by means of a grid search and sequential quadratic programming optimization algorithms. In addition, these studies revealed that several sets of optimization parameters deliver very similar fits to the experimental PELDOR data due to the inversion symmetry of the spin Hamiltonian. This symmetry makes PELDOR time traces insensitive to a 180° rotation of spin centers around one of their *g* principal axes.

A model-free approach was proposed by Marko *et al.*[103] Their fitting algorithm reconstructs an experimental dataset by searching for an optimal combination of pre-simulated PELDOR time traces of all relative orientations and inter-spin distances in the experimentally accessible range. Good fits to experimental PELDOR data were obtained on a model bis-nitroxide systems investigated. The authors pointed out that the solutions were not unique, which might be a consequence of the insufficient orientation selectivity achieved at the X-band frequencies.

Despite the existence of several approaches outlined above, none of them provides a general way to analyze orientation selective PELDOR data independently from the type of a spin system. The lack of such a generalized approach significantly complicates the extraction of distance information from PELDOR time traces acquired on rigid spin centers.

1.4.8 Applications of PELDOR

PELDOR-based distance measurements have evolved into a growing field in biological and material sciences. The number of related publications, especially those involving biologically relevant systems, has rapidly increased in the past years: searching for the keywords “PELDOR” and “DEER” by the Google Scholar engine (scholar.google.com) yields 200 publications in the time period of 1981-2010 and 563 publications in 2010-2016. To systemize all PELDOR studies, two classifications are commonly used. One of them distinguishes PELDOR measurements on different types of molecular systems, e.g., on model compounds,[104] organic polymers,[105] soluble and membrane proteins,[23,106] and

oligonucleotides.[107] Another classification is based on the different types of paramagnetic centers used for PELDOR. These include artificially introduced spins labels, such as nitroxide,[23] gadolinium,[32] and trityl[34] labels, as well as naturally occurring spin centers, such as organic co-factors[97] and metal ions[14]. Furthermore, PELDOR measurements on molecules containing two or more different spin centers attract an increasing interest.[41] Besides that, the progress in the synthesis of spin labels with desired properties allowed recently application of PELDOR to biomolecules not only in frozen solution, but also in viscous liquid solution,[35,36] and even within cells[108,109].

In the present study, PELDOR-based distance measurements between metal ions and nitroxides are in focus. An overview of existing PELDOR studies on such spin pairs is given in the next section.

1.4.9 Applications of PELDOR to the pair metal ion-nitroxide

So far, PELDOR-based distance measurements between paramagnetic metal ions and nitroxides have mainly been carried out on model compounds, demonstrating the feasibility of such measurements and setting the stage for real applications. A metal ion that is most frequently used in PELDOR studies is Cu^{2+}. The spin center of the Cu^{2+} ion has relatively long relaxation times and a moderate *g*-anisotropy, which makes it a suitable probe for PELDOR experiments. The first Cu^{2+}-nitroxide PELDOR experiment was performed by Narr *et al.* on a rigid model compound containing one Cu^{2+} and two nitroxides.[110] To achieve a large modulation depth, the PELDOR time traces were acquired by setting the pump pulse to the maximum of the nitroxide spectrum and observing the Cu^{2+} signal. The recorded PELDOR time traces displayed a significant extent of orientation selection in the copper spectrum, which however was not analysed in that work. Instead, the Cu-nitroxide distance was only estimated from the experimental dipolar spectra. An explicit treatment of orientation selectivity was done for the PELDOR data of the nitroxide-labeled Cu^{2+} porphyrin by Bode *et al.*[98,111] The simulations were based on the known experimental and spin Hamiltonian parameters and make use of a geometric model as employed for structurally similar bis-nitroxides and spin density parameters as obtained from density functional theory calculations. In addition to the distance, the authors were able to determine the exchange coupling constant between two spin centers. A similar analysis has been recently reported by Bowen *et al.* for a porphyrin-based Cu^{2+}-nitroxide model system with limited flexibility.[100] An example of distance measurements between a Cu^{2+} ion and MTSSL in a small peptide is provided by the work of Yang *et al.*[112] The analysis of the PELDOR data was done using an

MD-based structural model of the peptide and assuming some dynamics of the peptide and MTSSL. In another work from Yang *et al.*,[16] the PELDOR measurements on the Cu^{2+}-MTSSL pair have been used to narrow down the location of the Cu^{2+} ions in EcoRI-DNA complex.

Another biologically relevant metal ion, which was considered for PELDOR studies is Fe^{3+} in the low-spin state. Due to the significant *g*-anisotropy of the Fe^{3+} spin center, the application of PELDOR to the Fe^{3+}-nitroxide pair is more challenging as compared to the Cu^{2+}-nitroxide pair. In particular, broadband resonators are required to be able to apply the m.w. pulses at both the nitroxide spectrum and each *g*-component of the Fe^{3+} spectrum. Nevertheless, Lovett *et al.* demonstrated a successful PELDOR distance measurement between an MTSSL attached to palustrisredoxin reductase (PuR) and reduced (Fe^{3+}/Fe^{2+}) form of the [2Fe-2S] cluster in palustrisred oxin (Pux-B).[99] The detection pulses were positioned at the maximum of the nitroxide spectrum so that all orientations of the nitroxide with respect to the magnetic field are excited, and the pump pulse was applied at five different positions across the [2Fe-2S] EPR spectrum. This yielded PELDOR time traces with small modulation depths, but the SNR was good enough to let the authors thoroughly analyze these time traces and extract an accurate distance from them. Importantly, using this distance constraint enabled a well-defined model of the PuR/Pux-B complex to be established. In another work from Ezhevskaya *et al.*,[113] the distance between the Fe^{3+} ion and MTSSL have been measured in haem-containing and spin-labeled human neuroglobin. Due to very low SNR of the acquired PELDOR time traces, the orientation effects were not analyzed in that work. Using the W-band EPR spectrometer HiPER and composite pulses, Motion *et al.* could improve the sensitivity of the latter measurements by a factor 30.[81] However, also here the authors did not consider orientation selectivity effects, explaining this by the fact that the application of the m.w. pulses to the g_{yy} components of the iron and the nitroxide spectra leads a broad range of dipolar angles being excited.

Since naturally occurring Mg^{2+} ions can be often exchanged for Mn^{2+} ions, PELDOR measurements on Mn^{2+} ions are of high interest. The difficulty of PELDOR experiments on these high-spin ions ($S = 5/2$) is related to relatively short relaxation times and the presence the multiple electronic transitions with different transition moments. Moreover, the PELDOR experiments have to be usually performed at high frequencies, such as Q-band frequencies (~ 34 GHz) or higher, in order to ensure that the high-field approximation ($H_{ZFS} \ll H_{EZ}$ in Equation 1.4.1) is valid and that both spins are quantized along the direction of the static magnetic field. The first Mn^{2+}-nitroxide distance measurements by Q-/G-band PELDOR were

accomplished by Akhmetzyanov *et al.* for a Mn^{2+}-terpyridine complex.[114] These measurements revealed that orientation selectivity can be neglected due to the flexibility of the nitroxide and a broad statistical distribution of the zero-field splitting parameters of Mn^{2+}. A similar result was also obtained for the homoleptic bis-nitroxide complex of Mn^{2+} by Meyer *et al.*[115] Due to the low degree of ligand dissociation and two nitroxide spins per each molecule, the gain in the modulation depth of the PELDOR time traces of 250% as compared to the first work was reported. Recently, Kaminker *et al.* reported a first biological application of PELDOR-based Mn^{2+}-nitroxide distance measurements in the HP92 loop of ribosomal RNA.[116] In this study, the distances between the intrinsic Mn^{2+} ion and the spin labels 4-isocyanto-TEMPO attached to two different sites of RNA were explored. The obtained PELDOR time traces had low modulation depths and no oscillation, which led to very broad distance distributions. Nevertheless, using these distance constraints and the results of ESEEM/ENDOR experiments, the authors could propose a possible location of the Mn^{2+} ion within the structure of the HP92 loop.

Additionally, the distance between a high-spin Gd^{3+} ion ($S = 7/2$) and a nitroxide was explored by PELDOR in several model systems[117,118] and biomolecules[119,120]. Note that Gd^{3+} is not a naturally occurring metal ion in biology and is used only as an EPR-active center of spin labels. Many Gd^{3+} centers possess a large distribution of zero-field splitting parameters (***D*** in Equation 1.4.1), thus significantly reducing effects of orientation selectivity in PELDOR experiments. The values of zero-field splitting parameters of Gd^{3+} are larger than those of Mn^{2+}. Therefore the PELDOR experiments on Gd^{3+} are usually performed at W-band frequencies (~ 95 GHz), in order to ensure realization of the high-field approximation ($H_{ZFS} \ll H_{EZ}$ in Equation 1.4.1). Moreover, the central transition of the Gd^{3+} spectrum narrows considerably as the frequency increases, which leads to a considerable sensitivity increase for the high-frequency PELDOR measurements.

1.4.10 RIDME

The pulse sequence for the 5-pulse RIDME experiment is shown in Figure 1.4.4f. All m.w. pulses are applied at one frequency, which is set in resonance with one of the coupled spins. In analogy to the PELDOR description, this spin will be denoted here as A spin. The second spin, which is not affected by m.w. pulses, will be called B spin. The first π/2- and π-pulses of the sequence create a primary spin echo (Hahn echo). This primary echo is then refocused by the third and fourth π/2-pulses that are applied at a time *t* after the primary echo and create two additional echoes, a stimulated echo at a time *t* after the fourth pulse and a virtual echo at

a time t before the fourth pulse. Strictly speaking, a virtual echo is not a real echo, because it describes a spin ensemble whose transversal magnetization is defocusing during the time t. However, applying an additional π-pulse at a time (τ_2 - t) after the fourth pulse, the virtual echo can be refocused into a real echo appearing at a time τ_2 after this π-pulse. The flip of the B spins occurs due to spontaneous relaxation events in the time interval T between the third and the fourth pulses. The highest flip probability is achieved when T is on the order of the T_1 relaxation time of the B spins. Moreover, T should be much longer than other inter-pulse intervals in the sequence, so that the T_1 relaxation of the B spins can be neglected in these intervals. In analogy with the PELDOR experiment, the flip of the B spins causes a shift of the Larmor frequency of the A spins after the fourth pulse,

$$\omega_{AB}^{kl} = \omega_{dd}\left(1 - 3\cos^2\theta\right)\left(m_k - m_l\right) = \omega_{dd}\left(1 - 3\cos^2\theta\right)\Delta m_{kl} \tag{1.4.26}$$

where ω_{dd} and θ are defined by Eq. 1.4.6-1.4.8, m_k and m_l are two different projections of the $\boldsymbol{S}_z$ operator for the B spin. Due to this frequency shift, the precessing magnetization of the A spins accumulates an additional phase of $\omega_{AB}t$, which depends on the value of t. Thus, if the time interval t is linearly incremented in the experiment, the intensity of the refocused virtual echo will be modulated by dipolar frequencies. Such modulated echo signal is recorded in the RIDME experiment yielding a RIDME time trace.

Similar to the PELDOR time trace, the RIDME time trace can be written as a product of two components,

$$V_{RIDME}(t) = V_{signal}(t)V_{bckg}(t). \tag{1.4.27}$$

$V_{bckg}(t)$ describes a gradual decay of the echo signal due to intermolecular interactions of spin centers and spin diffusion. Due to the lack of an exact expression for $V_{bckg}(t)$, its shape is fitted empirically, usually as a stretched exponential function.[87,121] In addition, $V_{bckg}(t)$ may include electron spin echo envelope modulation (ESEEM) artifacts, which appear due to the hyperfine interaction of the A spins with surrounding nuclear spins with $I > 0$. Since the nuclear frequencies may overlap with the dipolar frequencies and thus complicate the data analysis, ESEEM has to be removed from RIDME time traces. Several approaches to do this are described in the literature.[87,93,122,123]

The second part to the RIDME time trace, $V_{signal}(t)$, describes the contribution to the echo signal which is caused by the dipole-dipole interaction between the A and B spins. In analogy to Equation 1.4.24 for $V_{intra}(t)$ of the PELDOR time trace, $V_{signal}(t)$ for N different conformers of a spin pair in a disordered powder sample can be written as

$$V_{signal}(t) = \sum_{i=1}^{N} \int_0^{2\pi} \int_0^{\pi} V_{0i}(\varphi,\psi) \left\{ 1 - \sum_k \sum_{l \neq k} \lambda_{kl} \left[1 - \cos\left(\omega_{AB}^{kl,i}(\varphi,\psi) t \right) \right] \right\} \sin\varphi d\varphi d\psi, \quad (1.4.28)$$

where V_{0i} is the A spin echo intensity at $t = 0$ for the i-th conformer, $\omega_{AB}^{kl,i}$ is the dipolar frequency of i-th conformer and is given by Equation 1.4.26, λ_{kl} are the modulation depth parameters, φ and ψ are two polar angles which describe the random orientation of the magnetic field. The modulation depth parameters λ_{kl} are defined as a product of two probabilities, the probability P_k of the B spin having the spin projection m_k, and the probability W_{kl} of the spin B changing its projection from m_k to m_l during the time interval T:

$$\lambda_{kl} = P_k W_{kl}. \quad (1.4.29)$$

If the correlation between the orientations of the A spins and the distance vectors can be neglected, Equation 1.4.28 can be simplified yielding

$$V_{signal}(t) = \langle V_0 \rangle \left\{ 1 - \sum_k \sum_{l \neq k} \lambda_{kl} \int_0^{\infty} D(r) \int_0^{\pi/2} \left[1 - \cos\left(\omega_{AB}^{kl}(r,\theta) t \right) \right] \cdot \sin\theta \, d\theta dr \right\}, \quad (1.4.30)$$

where $<V_0>$ is the average A spin echo intensity at $t = 0$, and $D(r)$ is the probability density function for the inter-spin distances. For the B spins with $S = 1/2$, Equation 1.4.30 can be further simplified:

$$V_{signal}(t) = \langle V_0 \rangle \left\{ 1 - \lambda \int_0^{\infty} D(r) \int_0^{\pi/2} \left[1 - \cos\left(\omega_{AB}(r,\theta) t \right) \right] \cdot \sin\theta \, d\theta dr \right\}. \quad (1.4.31)$$

Here, the dipolar frequency ω_{AB} is given by Equation 1.4.19, and the modulation depth parameter λ can be calculated as

$$\lambda = \frac{1}{2} \left[1 - \exp\left(-\frac{T}{T_1} \right) \right], \quad (1.4.32)$$

where T_1 is the longitudinal relaxation time of the B spins. Note that Equation 1.4.31 is almost identical to Equation 1.4.25. Thus, in the present case the distance distribution $D(r)$ can be extracted from the RIDME data using the same mathematical algorithms as used for the PELDOR data analysis, e.g. using Tikhonov regularization.

For B spins with $S > 1/2$, the transitions between different m_k and m_l yield

$$\left| m_k - m_l \right| = 0, 1, ..., (2S+1), \quad (1.4.33)$$

which leads to the modulation of the RIDME time trace with higher harmonics of ω_{AB}. In this case, the values of λ_{kl} for each transition have to be explicitly included into the analysis of the RIDME data.

When the correlation between the orientations of the A spins and the distance vectors is significant, the RIDME time trace will be orientation selective. The analysis of the orientation selective RIDME data has to be done using Equation 1.4.30 and it encounters the same difficulties as the analysis of orientation selective PELDOR data (see Section 1.4.6). Importantly, since the flip of the B spins in the RIDME experiment is independent of the orientation of these spins, only the detected A spin centers have to be flexible to completely avoid orientation selectivity in the RIDME experiment. This condition can be realized when e.g. the A spin corresponds to a flexible spin label.

1.4.11 Applications of RIDME

Compared to PELDOR, RIDME has significantly less applications so far. Most applications are nicely summarized in a recent review article from Astashkin.[124] Whereas the pioneering RIDME experiments were conducted on nitroxide biradicals,[56,86] more recent studies employed RIDME mostly for distance measurements between organic radicals and native metal centers.[93,122,123,125–129] Note that all of these studies were performed using the 3-pulse and 4-pulse RIDME sequences which suffer from dead time. The dead-time free 5-pulse RIDME sequence, which is used in this work, was proposed relatively recently and has only few applications in the literature. Before this thesis, 5-pulse RIDME was applied only to a nitroxide biradical and MTSSL-labeled cytochrome *f*. During and after the present work, the applications of 5-pulse RIDME were extended to two Gd^{3+}-based model compounds,[130,131] a homoleptic bis-nitroxide complex of Mn^{2+},[115] and the Gd^{3+}-labeled homodimeric protein ERp29[131]. Some of these studies which deal with distance measurements between metal ions and nitroxides are discussed in the next section.

1.4.12 Applications of RIDME to the pair metal ion-nitroxide

Up to date, distance measurements between a metal ion and a nitroxide by RIDME were explored in two works. In the first work, Milikisyants *et al.* tested the 5-pulse RIDME sequence on the Fe^{3+}-MTSSL spin pair in MTSSL-labeled cytochrome *f*.[87] The RIDME signal was detected on the slower-relaxing nitroxide spins, whereas the fast-relaxing Fe^{3+} spins were flipped due to spontaneous relaxation. The flip probability of Fe^{3+} was maximized by optimizing the inter-pulse interval in which the relaxation events happen. The acquired RIDME time trace displayed a modulation depth of 24%, but no oscillation. As a consequence, the extracted Fe^{3+}-MTSSL distance distribution was almost 1.5 nm wide. In the recent study by Meyer *et al.*,[115] the RIDME measurements were applied to the Mn^{2+}-

nitroxide model compound. Again, the nitroxide echo signal was detected, and the Mn^{2+} spins were flipped due to spontaneous relaxation. The RIDME time traces displayed slight orientation selectivity in dependence of the detection positions in the nitroxide spectrum. To suppress the orientation selectivity effects, the authors averaged the RIDME time traces over the different detection positions across the nitroxide spectrum. Since the Mn^{2+} has a spin of 5/2, the RIDME time traces were found to contain higher harmonics of ω_{dd}, which correspond to Δm_{kl} = 1, 2, and 3 in Equation 1.4.26. This feature was first reported by Razzaghi *et al.* for the Gd^{3+} ions with $S = 7/2$.[130] Due to higher dipolar harmonics, the analysis of the RIDME traces by DeerAnalysis yielded two ghost distance peaks corresponding to $\Delta m_{kl} = 2$ and $\Delta m_{kl} = 3$.

1.5 Trilateration

Trilateration is a method of determining the absolute or relative location of an object by means of the distances measured between this object and a number of reference points with defined location. It is commonly used in surveying and navigation, including global positioning systems (GPS). Applications of trilateration in biology are also known.[132,133]

The basic principle of trilateration is demonstrated in Figure 1.5.1. Each measured distance r_i between an object to be localized and a reference point with coordinates $\boldsymbol{p_i}$ restricts the possible location of an object to a surface of a sphere which is centered at $\boldsymbol{p_i}$ and has a radius of r_i. In case of several measured distances for different reference points, the location of the object is defined by the area where the corresponding spheres intersect each other. In the ideal

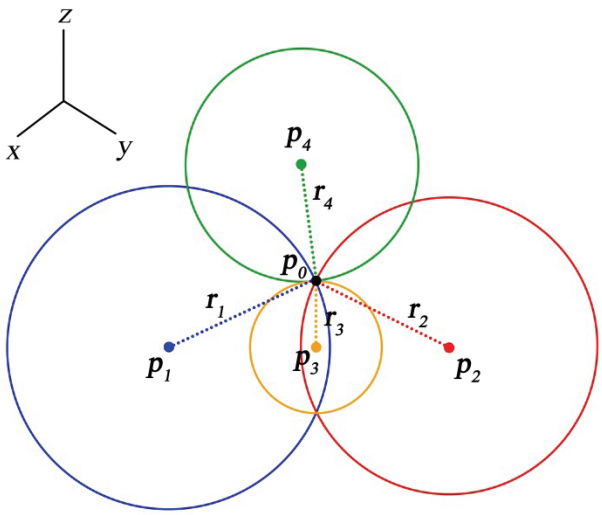


Figure 1.5.1. The trilateration principle. The location of the object, $\boldsymbol{p_0}$, is determined by means of the distances $\boldsymbol{r_i}$ measured between this object and reference points with coordinates $\boldsymbol{p_i}$ (i = 1, 2, 3, 4). The position of the object is given by the point in which all four spheres with the centers at $\boldsymbol{p_i}$ and the radii of $\boldsymbol{r_i}$ intersect each other.

case, when all coordinates $\boldsymbol{p_i}$ and distances r_i can be determined precisely, 4 reference points and 4 distances are sufficient to determine the exact position of the object. In practice, however, $\boldsymbol{p_i}$ and r_i are measured with some error and therefore the use of more than 4 distance constraints is advantageous. For N distance constraints the trilateration problem is described by the following system of equations:

$$(\boldsymbol{p_i} - \boldsymbol{p_0})^2 - r_i^2 = 0, \quad i = \overline{1, N}, \tag{1.5.1}$$

where $\boldsymbol{p_0} = [p_{0x}, p_{0y}, p_{0z}]$ are the unknown coordinates of the object to be localized. To solve this system of equations, several mathematical algorithms are available in the literature.[134] One of them is the Levenberg-Marquardt method.[134] The goal of this method is minimization of the χ^2 merit function:

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{r_i - \sqrt{(\boldsymbol{p_i} - \boldsymbol{p_0})^2}}{\sigma_i} \right]^2, \tag{1.5.2}$$

where σ_i is the standard error of the measured distance r_i. Given an initial guess for $\boldsymbol{p_0}$, the minimization of χ^2 proceeds through the following iterations:

1) Calculation of $\chi^2(\boldsymbol{p_0})$.
2) Calculation of the matrices

$$\beta_k = -\frac{1}{2}\frac{\partial \chi^2(\boldsymbol{p_0})}{\partial p_{0k}}, \alpha_{kl} = \frac{1}{2}\frac{\partial^2 \chi^2(\boldsymbol{p_0})}{\partial p_{0k} \partial p_{0l}} + \lambda \delta_{kl}, \tag{1.5.3}$$

   where k and l are x, y, or z, δ_{kl} is the Kronecker delta, and λ is an adjustable damping parameter, which is initially set to 0.001.
3) Determination of the increment for $\boldsymbol{p_0}$, $\delta\boldsymbol{p_0}$, by solving the system of linear equations

$$\sum_{l=1}^{3} \alpha_{kl} \delta p_{0l} = \beta_k. \tag{1.5.4}$$

4) Evaluation of $\chi^2(\boldsymbol{p_0} + \delta\boldsymbol{p_0})$. If $\chi^2(\boldsymbol{p_0} + \delta\boldsymbol{p_0}) \geq \chi^2(\boldsymbol{p_0})$, λ is increased by a factor of 10 and the calculations continue with step 2. If $\chi^2(\boldsymbol{p_0} + \delta\boldsymbol{p_0}) < \chi^2(\boldsymbol{p_0})$, the trial solution for $\boldsymbol{p_0}$ is replaced by $(\boldsymbol{p_0} + \delta\boldsymbol{p_0})$, λ is decreased by a factor of 10, and the calculations continue with step 2.

The optimization is finished when $\chi^2(\boldsymbol{p_0} + \delta\boldsymbol{p_0})$ decreases with respect to $\chi^2(\boldsymbol{p_0})$ by a factor of less than 10^{-3}. The standard error of the calculated $\boldsymbol{p_0}$ coordinates can be estimated as

$$\sigma(p_{0k}) = \sqrt{(\alpha^{-1})_{kk}}, \tag{1.5.5}$$

where α^{-1} is the inverse α_{kl} matrix, which is calculated from Equation 1.5.3 using the optimized coordinates $\boldsymbol{p_0}$ and $\lambda = 0$.

To conclude, trilateration represents a method to localize objects in three-dimensional space. The main requirements of this method are the presence of 4 or more reference points with a well-defined location in the coordinate system of the object and the possibility to measure the distances between the reference points and the object of interest.

1.6 Aim of the work

The aim of the work is the development and testing of an EPR-based metal ion localization method for biomolecules. The method is supposed to be based on the trilateration principle, according to which the position of a metal ion in a biomolecular structure can be determined from the distances measured between the ion itself and a number of spin labels attached to the biomolecule using SDSL. To solve the trilateration problem, a new computer software will be developed. The evaluation of the EPR-based metal ion localization method will be done on the Cu^{2+} ion in the protein azurin. Since the Cu^{2+} site of azurin is well characterized by X-ray crystallography, the precision of the EPR-based method will be determined, and the factors that affect this precision will be evaluated.

The use of the trilateration principle assigns an important role to EPR-based distance measurements between metal ions and spin labels. Thus, a major goal of this study is the determination of efficient EPR methods for the distance measurements involving metal ions. As discussed in Section 1.7, PELDOR measurements on metal ions usually yield orientation selective time traces, whose analysis is very challenging. Therefore, a general approach to analyze the orientation-selective PELDOR data will be developed. In addition, the RIDME technique will be evaluated as a possible alternative to PELDOR. To do this, the performance of PELDOR and RIDME will be compared for the Fe^{3+}-nitroxide spin pair of the MTSSL-labeled protein cytochrome P450cam.

Another important aspect of the trilateration is related to the determination of the spin labels' positions in biomolecular structures. Since most of the spin labels have some intrinsic flexibility, the determined positions depend on the conformational ensembles of spin labels. Therefore, one of the goals of this study is an investigation of the conformational ensembles of spin labels in biomolecules. First, the conformations of MTSSL in azurin mutants will be explored by PELDOR and X-ray crystallography. Second, EPR and DFT simulations will be used to characterize the intrinsic flexibility of a nitroxide-labeled uracil nucleobase in RNA.

2 Results and Discussion

2.1 Development of an EPR-based metal ion localization method

The EPR-based metal ion localization method in this work makes use of the trilateration principle. According to this principle, the location of a paramagnetic metal ion in a biomolecule can be deduced from distances measured between this ion and a number of spin labels attached to the surface of the biomolecule by site-directed spin labelling. The attached spin labels act as reference points in the molecular coordinate system of the biomolecule. The trilateration problem defined this way was accomplished in the following steps:

1. Synthesis and spin labeling of the biomolecule.
2. Measuring the distances between the metal ion and the attached spin labels by means of EPR.
3. Choosing the structural model of the biomolecule and the coordinate system associated with it.
4. Determining the coordinates of the attached spin labels in the coordinate system of the biomolecule.
5. Solving the trilateration problem using the determined distances and coordinates of the spin labels.

Taking into account extensive experience of biologists in synthesis and modification of various biomolecules as well as a broad arsenal of spin labeling techniques, step 1 can be readily achieved for the majority of biomolecules.

Step 2 is related to distance measurements between the metal ion and the attached spin labels by means of pulsed EPR. Among all pulsed EPR techniques, PELDOR is the most commonly used one. However, the analysis of experimental PELDOR time traces acquired on metal centers is often complicated by orientation selectivity effects. To enable the extraction of distances from orientation-selective PELDOR data, a method based on a simplified model of a spin system was developed in this work. The description of the method is given in Section 2.2. Additionally, an alternative to the PELDOR technique, RIDME, was evaluated for the distance measurements on a metal ion/nitroxide spin pair. A detailed comparison of PELDOR and RIDME for the spin-labeled mutant of cytochrome P450cam is described in Section 2.6.

The next step of the trilateration is choosing the structural model of the biomolecule. This step depends on the amount of structural information available for the particular biomolecule. In the best case, the crystal structure of the metal-free form of the biomolecule can be used as a

starting structure for the trilateration. Another possibility is to use a homology model of the biomolecule. The programs that allow creating such models are available for proteins[135] and oligonucleotides[136].

As soon as the model of the biomolecule has been chosen, the coordinates of the attached spin labels were determined. Since the EPR-based distance measurements are performed between the spin centers of the spin labels, the coordinates of the spin centers are actually of interest. The determination of these coordinates is complicated by the fact that the commonly used spin labels, such as MTSSL, contain a flexible linker between one end, where they bind to the biomolecule, and another end, where the spin center is. Thus, the coordinates of the spin centers have to be averaged over different conformations of the spin labels in an ensemble of identical biomolecules. To do this, one needs to know the conformational ensemble for each spin label. In this work, such conformational ensembles were modeled by means of computer simulations called in the literature as *in silico* spin labeling. Two commonly used *in silico* methods are based on the accessible volume approach[137] and the rotamer approach[138]. The former is used by the program mtsslWizard,[139] the latter one by MMM[138] and PRONOX[140]. In this work, the mtsslWizard and MMM programs were mostly used to model the conformational ensembles of spin labels and to calculate the average coordinates for their spin centers. Besides the coordinates of the spin labels, these programs provide an important opportunity to plan the positions of the spin labeled sites. Such planning helps to achieve two important goals: 1) the chosen labeling sites should be well exposed in order to promote an efficient spin labeling, 2) the labeling sites should be approximately evenly distributed on the biomolecular surface to avoid the undesirable situation when all labels are concentrated on one side with respect to the metal ion site.

The last step of the trilateration is the calculation of the metal ion's position from the determined earlier distances and coordinates of the spin labels. The calculations were done in accordance to the algorithm discussed in Section 1.5. This algorithm was implemented into the program mtsslTrilaterate that is described in Section 2.3.

All five steps described above were realized for the Cu^{2+}-containing protein azurin that was used as a test biomolecule for the EPR-based metal ion localization method. The details of the trilateration procedure and the precision of the obtained Cu^{2+} location are described in Section 2.4.

2.2 Geometric model based analysis of orientation selective PELDOR data [P1]

In order to enable the analysis of orientation selective PELDOR data without any preliminary knowledge about the relative orientation of the spin centers, a fitting approach was introduced. This approach uses a simplified geometric model to describe the relative position and orientation of two spin centers involved in the PELDOR experiment. The geometric model is described by six parameters: the length of the distance vector, two spherical angles which describe the orientation of one of the spin centers with respect to the distance vector, and three Euler angles which describe the relative orientations of two spin centers. Each of these parameters can have either uniform or normal distribution, which are described by only two values, a mean value and a distribution width. The parameters of the model are optimized by means of a genetic algorithm to provide the best agreement between the experimental PELDOR time traces and the corresponding time traces simulated for the model. The simulation of PELDOR time traces requires also the spectroscopic characteristics of spin centers, e.g., $\boldsymbol{g}$ and hyperfine tensors, as well as the experimental settings, e.g. m.w. frequencies and pulse lengths. To perform the optimization within a reasonable time, two simplifying assumptions are applied to the model: 1) the spin centers are considered as single-point objects, and 2) the correlation between individual geometric parameters is neglected. Further details of the optimization procedure are given in Appendix P1.

This fitting approach was implemented in a freely available program called PeldorFit. Its manual can be downloaded from https://github.com/dinarabdullin/PeldorFit2015. The program was benchmarked on several orientation-selective PELDOR data sets from the literature, revealing a good agreement between the optimized geometric models and the previously reported models for the spin systems studied. Particular attention was put on the uniqueness and precision of the PeldorFit-derived models. It was shown that the inter-spin distance distribution could be unambiguously derived from the PELDOR time traces, whereas the relative orientation of the spin centers is determined up to several symmetry-related sets of angles. The precision of the optimized parameters was evaluated by recording the goodness-of-fit as a function of each parameter. This showed that the number of well-defined angular parameters of the model is strongly dependent on the degree of orientation selectivity. In contrast, the precision of the distance parameters was higher than 0.1 nm due to the high quality of the analyzed PELDOR data. Further applications of the PeldorFit program are given in Appendices P4 and P7. All tests reveal that the PeldorFit program provides a robust tool to extract the distance distributions from the orientation-selective PELDOR data and, therefore, can be applied for the PELDOR data acquired on metal centers.

2.3 *In silico* spin labeling and trilateration [P2, P3]

As mentioned in Section 2.1, the program mtsslWizard was extensively used in this work to construct spin-labeled models of biomolecules and to estimate average coordinates of attached spin labels in the coordinate system of biomolecules. The initial version of the program, which was introduced by Hagelueken *et al.* in 2012,[139] was made for prediction of conformational ensembles of MTSSL. Here, the program was extended to other spin labels and supplemented by additional tools for planning of spin-labeling sites. Moreover, it was combined with another two programs called mtsslTrilaterate and mtsslDock in a software package mtsslSuite. mtsslTrilaterate and mtsslDock were designed to use the mtsslWizard-generated models of spin labels for the localization of spin centers via trilateration and the reconstruction of macromolecular complexes via distance constrained rigid-body docking, respectively. In this thesis, only mtsslTrilaterate was employed. This program provides a graphical user interface for solving the trilateration problem by means of the mathematical algorithm described in Section 1.5. Like mtsslWizard, mtsslTrilaterate works as a plugin to a free available molecular visualization program PyMOL (www.pymol.org) that provides a nice graphical platform to display the starting structure of the biomolecule, the models of spin labels, and the calculated position of the metal ion. The performance of the mtsslTrilaterate program was confirmed by reproducing the trilateration results for the TEMPO-labeled lipid in soybean seed lipoxygenase-1 published by Gaffney *et al.*[15]. Another application of mtsslTrilaterate is described in Section 2.4. Furthermore, Evans *et al.* have recently used the program to calculate the location of the Cu^{2+} ion in the N-terminal octarepeat domain of the protein prion.[141]

2.4 Application of the EPR-based metal ion localization method to the Cu^{2+} ion in azurin [P4]

This chapter focuses on testing the EPR-based metal ion localization method introduced in Section 2.1. The Cu^{2+} ion in the protein azurin was chosen as a test trilateration target because 1) the synthesis and mutation of this protein are well established, 2) the crystal structures of both the metal-free and metal-bound forms of azurin are available, 3) the Cu^{2+} ion is EPR spectroscopically well characterized, 4) PELDOR measurements involving Cu^{2+} ions are most established among all metal ions. To perform the trilateration, six single cysteine mutants of azurin were synthesized and labeled with MTSSL. The formation of the R1 side chains after the labeling was confirmed by CW-EPR. The average coordinates of the N-O spin centers of the R1 side chains within the metal-free structure of azurin were estimated using the programs mtsslWizard and MMM. The distance measurements between the Cu^{2+} ion and the R1 side chains were performed by means of the PELDOR technique. To account for orientation selectivity, four PELDOR time traces with four different detection positions across the copper spectrum were acquired for each mutant. These time traces were further analyzed by means of the program PeldorFit (see Section 2.2), which yielded the Cu^{2+}-R1 distance distributions. Using these distance distributions and the average N-O coordinates of the R1 side chains, the location of the Cu^{2+} ion in the azurin structure was calculated by means of the program mtsslTrilaterate (see Section 2.3). The obtained solution was compared with the crystallographic data of the copper-bound azurin, which revealed that the calculated location of the Cu^{2+} ion is within the actual binding site and deviates from the corresponding crystallographic location by only 0.26 nm. To determine the origins of this deviation, all factors which affect the trilateration result were evaluated. These include the accuracy of the starting protein structure, the number of distance constraints, the precision of the distance constraints, the precision of the average N-O coordinates, and the spin density distribution of the Cu^{2+} ion. The first three factors were shown to have an only minor contribution to the deviation obtained. In contrast, taking into account the spin density distribution of the Cu^{2+} ion, the deviation could be reduced down to 0.19 nm. The remaining deviation was assigned to the errors of the average N-O coordinates. These errors are due to the fact that existing *in silico* methods cannot predict conformational ensembles of the R1 side chains exactly. This problem is considered in more detail in the next section.

2.5 Determination of MTSSL conformers in azurin [P5]

The conformational ensembles of the R1 side chains in the six azurin mutants were investigated using orientation-selective PELDOR and X-ray crystallography. Although the PELDOR data alone does not provide information about the structure of the R1 side chains, it can be still used to constrain the positions and orientations of the corresponding N-O spin centers relative to the rigid Cu^{2+} spin center. Using the known position and orientation of the Cu^{2+} spin center within the azurin structure, the PELDOR constraints were used to localize the N-O groups of the R1 side chains in the azurin structure. The localization was done in two steps. In the first step, the conformational ensembles of the R1 side chains were generated by means of the programs mtsslWizard and MMM. These ensembles yielded the Cu^{2+}-R1 distance distributions that are prominently broader than the ones obtained from the PELDOR measurements, thus revealing that the applied *in silico* methods overestimate the conformational space sampled by the R1 side chains in azurin. The mean values of the calculated and PELDOR-derived distance distributions deviated in average by 0.15 nm. In the second step, the R1 side chain conformers from the initial ensembles were sorted out based on the fitness to the PELDOR time traces. This task required the use of an optimization algorithm, in order to ensure that the selected conformers deliver the best fitness to the PELDOR time traces. The R1 side chain conformers obtained after such optimization were compared with the crystal structures of the spin-labeled azurin mutants T21R1 and T30R1. The comparison revealed significant similarities in the positions and orientations of the corresponding N-O groups, whereas the structure of the R1's flexible linker differed between the X-ray and PELDOR-derived conformers. The latter result can be explained by the fact that the PELDOR data is insensitive to the rotatable dihedral angles of the R1 side chain. Nevertheless, the similarity of the N-O positions determined by PELDOR and X-ray crystallography allows us to assume that the R1 side chain adopts very similar conformations in crystal phase and in frozen solution.

2.6 Comparison of RIDME and PELDOR [P6]

One problem that significantly complicates PELDOR measurements on metal ions is that spectral widths of metal ions are much larger than the exitation window of commonly used m.w. pulses of 8-32 ns. As a consequence, an application of the m.w. pulses to an EPR spectrum of a metal ion results in selective excitation of only a small part of metal ions in a sample. Usually, this leads to a low signal-to-noise ratio (SNR) and orientation selectivity for the PELDOR time traces. Furthermore, spin couples which consist of a nitroxide spin label and a metal center with largely different ***g*** values are difficult to study with the double-frequency PELDOR technique due to the limited bandwidths of m.w. resonators and m.w. amplifiers. All these difficulties can be tolerated to some extent for the Cu^{2+} ion, which displays a moderate ***g***-anisotropy and relatively long relaxation times. However, for other metal centers, such as Fe^{3+}, the problems outlined above arise at full strength. As a possible solution, the use of an alternative pulsed EPR technique called RIDME was considered. This technique is well suited for spin couples that consist of a slow-relaxing and a fast-relaxing spin, which fits well to the case of a nitroxide spin label and a metal center. Another important feature of the RIDME experiment is that the fast-relaxing spins (e.g., the spins of metal centers) are flipped by non-selective spontaneous relaxation rather than by m.w. pulses as it is in the PELDOR experiment. This removes the need in extremely broadband m.w. resonators and can allow avoiding orientation selectivity. To explore these advantages of RIDME in more detail, the comparison of RIDME and PELDOR for distance measurements between the low-spin Fe^{3+} ion and MTSSL in a MTSSL-labeled mutant of the heme-containing protein cytochrome P450cam was performed. It was shown that, in contrast to the PELDOR time traces, the RIDME time traces are free from orientation selectivity effect and have 7 times higher SNR. This ensured an easy and robust conversion of the RIDME time traces into the distance distributions, whereas only a rough estimate of the distance could be obtained for PELDOR time traces. Additionally, special emphasis was put on the optimization of the dead-time free 5-pulse RIDME experiment, whose description in literature is limited. Thus, Appendix P2 summarizes recipes how to optimize the RIDME pulse sequence and how to get rid of unwanted ESEEM modulation in the RIDME time traces. The optimal time T for the inversion of the Fe^{3+} spin was determined from the temperature-dependent inversion recovery measurements. ESEEM was removed from the RIDME time traces by dividing two RIDME signals acquired with two different T value or at two different temperatures, as well as by using higher m.w. frequencies and selective m.w. pulses.

2.7 Extension of the EPR-based metal ion localization method to nucleic acids [P7]

One of the potentially important applications of the EPR-based metal ion localization method is related to ribozymes. Many ribozymes, such as hammerhead and Diels-Alder ribozymes, contain Mg^{2+} ions which can be readily substituted by paramagnetic Mn^{2+} ions to allow for EPR measurements on these systems.[142] The procedure of EPR-based distance measurements between Mn^{2+} ions and nitroxides has been recently reported by several research groups.[114,116] Besides the distance measurements, spin labels with well characterized structure and dynamics are required for the trilateration. In this work, the structure and dynamics of the nitroxide-labeled uracil nucleobase (**dU**) in RNA was investigated by means of EPR spectroscopy and DFT calculations. The CW-EPR spectra of the **dU**-containing double-stranded RNA revealed that **dU** is flexible to some extent at room temperature. However, the PELDOR measurements on the same RNA duplex yielded highly orientation selective time traces, and the derived distance distribution had the same width as the distance distribution measured on an identical RNA duplex with the rigid **Çm** labels instead of **dU**. To interpret this observation it was proposed that **dU** adopts a limited number of energetically favorable conformations upon lowering the temperature down to the freezing temperature of a solvent. The DFT simulations showed that these energetically favorable conformations would correspond to an almost planar geometry of the spin-labeled nucleobase. Based on this information, the structural model of **dU** was developed and implemented in the program mtsslWizard. The use of this model to predict the distances between two **dU**s in the RNA duplex provided a reasonable agreement with the PELDOR-derived distance distribution. The difference of the corresponding mean distances was only 0.1 nm. Note that the prediction of the distance distribution was done using an idealized structure of RNA from the program 3DNA and not its crystal structure. Therefore, some error in the RNA structure might also influence the obtained result.

3 Summary and Outlook

In this thesis, an EPR method for localization of metal ions in biomolecules was developed and successfully tested on the Cu^{2+} ion in the protein azurin. The location of metal ions was determined through trilateration, which required site-directed spin labeling of several sites in a biomolecule, followed by the distance measurements between the labeled sites and a metal ion. The applicability of the method depends on availability of the structural model of the biomolecule, as well as on the ability to determine the coordinates of the spin labels within this model and to measure the distances between the spin labels and the metal ion. Methods for obtaining these data were investigated in this work. The crystal structure of the metal-free form of a biomolecule and the homology model of a biomolecule were shown to be suitable starting structural models. The coordinates of the spin labels were obtained using computer programs mtsslWizard and MMM, which model spin labels into a structural model of a biomolecule. The distances between the metal ion and the spin labels were measured by means of the pulsed EPR techniques PELDOR and RIDME. The procedure of such distance measurements was explored on two biologically relevant metal centers, Cu^{2+} and Fe^{3+}. As expected, the PELDOR data acquired for both metal ions displayed significant orientation selectivity. Up to date, the extraction of distance distributions from the orientation-selective PELDOR data is very challenging, because it requires the knowledge of relative orientations of spin centers involved in the PELDOR experiment. To enable the extraction of distances from orientation-selective PELDOR data without any preliminary knowledge about these orientations, a method based on a simplified model of a spin system was developed and realized as a program PeldorFit. The extensive tests of the program on a number of PELDOR data sets from the literature and from present work confirmed its robustness and accuracy in determination of distance distributions. For example, PeldorFit allowed for extraction of six Cu^{2+}-nitroxide distance distributions from the orientation-selective PELDOR data of MTSSL-labeled azurin mutants. However, the application of PELDOR to metal centers with spectral widths larger than that of the Cu^{2+} ion was shown to be complicated not only by the need to analyze the orientation selectivity, but also by technical limitations, e.g., the limited bandwidths of m.w. resonators and amplifiers. For this case, the use of the alternative technique RIDME was considered. Since the dead-time free version of RIDME was introduced relatively recently and had only few applications in the literature, the optimization of the RIDME experiment for the Fe^{3+}-nitroxide spin pair was accomplished. The results of the RIDME measurements revealed that this technique allows avoiding orientation selectivity.

In addition, the RIDME signals had a 7 times higher signal-to-noise ratio as compared to the corresponding PELDOR signals. All this ensured an easy and robust conversion of the RIDME time traces into the distance distributions, whereas only a rough estimate of the distance could be obtained for PELDOR time traces. Thus, when used for the distance measurements between a metal ion and a spin label, the RIDME technique can be advantages than the PELDOR technique.

After the evaluation of the different methods for obtaining the input data of the trilateration problem, the EPR-based metal ion localization method was applied to the Cu^{2+} ion in azurin. The calculated location of the Cu^{2+} ion was within the same metal-binding site as identified earlier by X-ray crystallography. The precision of the trilateration was estimated at 0.26 nm. This value was shown to be affected by the spin density delocalization of the Cu^{2+} center, the precision of the MTSSL coordinates, the number and precision of the distance constrains, and the accuracy of the protein model. The main factor that limited the precision of the trilateration was the error of the MTSSL coordinates. This error stems from the imprecision of existing computational methods in predicting conformational ensembles of MTSSL. This problem was investigated for azurin mutants by comparing the conformational ensembles of MTSSL from the programs mtsslWizard and MMM with the corresponding ensembles derived from the experimental PELDOR and X-ray diffraction data. This study revealed that the computational methods significantly overestimate the conformational distributions of MTSSL, which results in the mean error of average N-O positions of 0.15 nm. Thus, the improvement of existing computational method is still required. Note that the effect of the unprecise MTSSL coordinates on the trilateration result can be partially alleviated by either increasing the number of the labeled sites/distance constraints or by constraining the positions of spin labels via additional distance measurements between the pairs of the spin labels that are used for the trilateration. Both approaches require, however, time-expensive synthesis of the singly- or doubly-labeled protein mutants, respectively. This difficulty might be overcome in future applying orthogonal spin labeling of a biomolecule with two different types of labels.

Even with the current precision, the EPR-based method could be a valuable tool for the localization of metal-binding sites in biomolecules. One potentially important application of the new method is related to ribozymes, in which the location of Mn^{2+} ions is frequently of interest. The trilateration of Mn^{2+} ions in ribozymes requires an established procedure to measure the distances between Mn^{2+} ions and spin labels, as well as spin labels with the known structure and dynamics in RNA. The first requirement was considered recently in

several publications, where the PELDOR and RIDME were applied to measure the distances between the Mn^{2+} ions and the nitroxides. As a possible candidate to fulfil the second requirement, a nitroxide spin label that attaches to a modified uridine nucleobase of RNA was investigated in this work. The conformational dynamics of this spin label was studied by means of EPR and DFT simulations. This study revealed that the nitroxide-labeled nucleobase has a similar rigidity to the spin label **Ç** that is commonly used for DNA. Based on the obtained information, the structural model of the nitroxide-labeled uracil nucleobase was proposed and implemented in the program mtsslWizard. The approbation of this model for the prediction of the experimental PELDOR distance between two such nucleobases revealed a reasonable accuracy of 0.1 nm. Thus, the investigated spin label can be applied in future for trilateration of the Mn^{2+} ions in hammerhead and Diels-Alder ribozymes.

References

[1] A. Messerschmidt, R. Huber, T. Poulos, K. Wieghardt, *Handbook of Metalloproteins Vol. 1-2*, John Willey & Sons, Chichester, **2001**.

[2] A. Messerschmidt, W. Bode, M. Cygler, *Handbook of Metalloproteins Vol. 3*, John Wiley & Sons, Chichester, **2004**.

[3] A. Messerschmidt, *Handbook of Metalloproteins Vol. 4-5*, John Wiley & Sons, Chichester, **2011**.

[4] N. V. Hud, *Nucleic Acid-Metal Ion Interactions*, Royal Society Of Chemistry, Cambridge, **2009**.

[5] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.

[6] A. Feig, O. Uhlenbeck, in *RNA World* (Eds.: R.F. Gesteland, T.R. Cech, J.F. Atkins), Cold Spring Harbor Laboratory Press, New York, **1999**, pp. 287–309.

[7] A. G. Tzakos, C. R. R. Grace, P. J. Lukavsky, R. Riek, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 319–342.

[8] M. R. Jensen, M. A. S. Hass, D. F. Hansen, J. J. Led, *Cell. Mol. Life Sci.* **2007**, *64*, 1085–1104.

[9] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *Dalton Trans.* **2008**, 3782–3790.

[10] J. R. Morrow, C. M. Andolina, in *Interplay between Met. Ions Nucleic Acids* (Eds.: A. Sigel, H. Sigel, O.R.K. Sigel), Springer Netherlands, Dordrecht, **2012**, pp. 171–199.

[11] J. W. Taraska, M. C. Puljung, N. B. Olivier, G. E. Flynn, W. N. Zagotta, *Nat. Methods* **2009**, *6*, 532–537.

[12] X. Yu, X. Wu, G. A. Bermejo, B. R. Brooks, J. W. Taraska, *Structure* **2013**, *21*, 9–19.

[13] S. Van Doorslaer, *Electron Paramagn. Reson.* **2008**, *21*, 162–183.

[14] D. Goldfarb, *Struct. Bond.* **2014**, *152*, 163–204.

[15] B. J. Gaffney, M. D. Bradshaw, S. D. Frausto, F. Wu, J. H. Freed, P. Borbat, *Biophys. J.* **2012**, *103*, 2134–2144.

[16] Z. Yang, M. R. Kurpiewski, M. Ji, J. E. Townsend, P. Mehta, L. Jen-Jacobson, S. Saxena, *Proc. Natl. Acad. Sci.* **2012**, *109*, E993–E1000.

[17] C. Remenyi, R. Reviakine, M. Kaupp, *J. Phys. Chem. B* **2007**, *111*, 8290–8304.

[18] M. Van Gastel, J. W. A. Coremans, H. Sommerdijk, M. C. Van Hemert, E. J. J. Groenen, *J. Am. Chem. Soc.* **2002**, *124*, 2035–2041.

[19] J. E. Roberts, J. F. Cline, V. Lum, H. B. Gray, H. Freeman, J. Peisach, B. Reinhammar, B. M. Hoffman, *J. Am. Chem. Soc.* **1984**, *106*, 5324–5330.

[20] R. Tsai, C. A. Yu, I. C. Gunsalus, J. Peisach, W. Blumberg, W. H. Orme-Johnson, H. Beinert, *Proc. Natl. Acad. Sci. U. S. A.* **1970**, *66*, 1157–63.

[21] C. Altenbach, T. Marti, H. G. Khorana, W. L. Hubbell, *Science* **1990**, *248*, 1088–1092.

[22] A. Doll, E. Bordignon, B. Joseph, R. Tschaggelar, G. Jeschke, *J. Magn. Reson.* **2012**, *222*, 34–43.

[23] G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.

[24] O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1–53.

[25] A. D. Milov, K. M. Salikhov, M. D. Shchirov, *Sov. Phys. Solid State* **1981**, *23*, 565–569.

[26] H. El Mkami, D. G. Norman, *Methods Enzymol.* **2015**, *564*, 125–152.

[27] L. J. Berliner, J. Grunwald, H. O. Hankovszky, K. Hideg, *Anal. Biochem.* **1982**, *119*, 450–455.

[28] G. Jeschke, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *72*, 42–60.

[29] E. G. Bagryanskaya, O. A. Krumkacheva, M. V. Fedin, S. R. A. Marque, *Methods Enzymol.* **2015**, *563*, 365–396.

[30] M. R. Fleissner, E. M. Brustad, T. Kalai, C. Altenbach, D. Cascio, F. B. Peters, K. Hideg, S. Peuker, P. G. Schultz, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 21637–21642.

[31] M. J. Schmidt, A. Fedoseev, D. Summerer, M. Drescher, *Methods Enzymol.* **2015**, *563*, 483–502.

[32] D. Goldfarb, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9685–9699.

[33] T. J. Reddy, T. Iwama, H. J. Halpern, V. H. Rawal, *J. Org. Chem.* **2002**, *67*, 4635–4639.

[34] G. W. Reginsson, N. C. Kunjir, S. T. Sigurdsson, O. Schiemann, *Chem. A Eur. J.* **2012**, *18*, 13580–13584.

[35] Z. Yang, Y. Liu, P. Borbat, J. L. Zweier, J. H. Freed, W. L. Hubbell, *J. Am. Chem. Soc.* **2012**, *134*, 9950–9952.

[36] G. Y. Shevelev, O. A. Krumkacheva, A. A. Lomzov, A. A. Kuzhelev, O. Y. Rogozhnikova, D. V. Trukhin, T. I. Troitskaya, V. M. Tormyshev, M. V. Fedin, D. V. Pyshnyi, et al., *J. Am. Chem. Soc.* **2014**, *136*, 9874–9877.

[37] D. Thonon, V. Jacques, J. F. Desreux, *Contrast Media Mol. Imaging* **2007**, *2*, 24–34.

[38] E. Matalon, T. Huber, G. Hagelueken, B. Graham, V. Frydman, A. Feintuch, G. Otting, D. Goldfarb, *Angew. Chemie Int. Ed.* **2013**, *52*, 11831–11834.

[39] Z. Yang, M. D. Bridges, C. J. López, O. Y. Rogozhnikova, D. V Trukhin, E. K. Brooks, V. Tormyshev, H. J. Halpern, W. L. Hubbell, *J. Magn. Reson.* **2016**, *269*, 50–54.

[40] A. Feintuch, G. Otting, D. Goldfarb, *Methods Enzymol.* **2015**, *563*, 415–457.

[41] M. Yulikov, *Electron Paramagn. Reson.* **2015**, *24*, 1–31.

[42] S. A. Shelke, S. T. Sigurdsson, *European J. Org. Chem.* **2012**, 2291–2301.

[43] J. P. Klare, H. J. Steinhoff, *Photosynth. Res.* **2009**, *102*, 377–390.

[44] N. Barhate, P. Cekan, A. P. Massey, S. T. Sigurdsson, *Angew. Chemie Int. Ed.* **2007**, *46*, 2655–2658.

[45] T. E. Edwards, S. T. Sigurdsson, *Nat. Protoc.* **2007**, *2*, 1954–1962.

[46] S. Saha, A. P. Jagtap, S. T. Sigurdsson, *Methods Enzymol.* **2015**, *563*, 397–414.

[47] P. Z. Qin, S. E. Butcher, J. Feigon, W. L. Hubbell, *Biochemistry* **2001**, *40*, 6929–6936.

[48] G. P. G. Grant, N. Boyd, D. Herschlag, P. Z. Qin, *J. Am. Chem. Soc.* **2009**, *131*, 3136–3137.

[49] A. Okamoto, T. Inasaki, I. Saito, *Bioorganic Med. Chem. Lett.* **2004**, *14*, 3415–3418.

[50] G. Sicoli, F. Wachowius, M. Bennati, C. Höbartner, *Angew. Chemie Int. Ed.* **2010**, *49*, 6443–6447.

[51] A. Ramos, G. Varani, *J. Am. Chem. Soc.* **1998**, *120*, 10992–10993.

[52] E. K. Zavoisky, *J. Phys. USSR* **1945**, *9*, 211–216.

[53] A. Abragam, B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions*, Oxford University Press, Oxford, **1970**.

[54] R. J. Blume, *Phys. Rev.* **1958**, *109*, 1867–1873.

[55] A. Schweiger, G. Jeschke, *Principles of Pulse Electron Paramagnetic Resonance*, Oxford University Press, New York, **2001**.

[56] L. V. Kulik, S. A. Dzuba, I. A. Grigoryev, Y. D. Tsvetkov, *Chem. Phys. Lett.* **2001**, *343*, 315–324.

[57] A. Bencini, D. Gatteschi, *Electron Paramagnetic Resonance of Exchange Coupled Systems*, Springer, Berlin, Heidelberg, **1990**.

[58] J. A. Weil, J. R. Bolton, *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*, John Wiley & Sons, Hoboken, **2007**.

[59] A. Lund, M. Shiotani, S. Shimada, *Principles and Applications of ESR Spectroscopy*, Springer, New York, **2011**.

[60] J. W. A. Coremans, O. G. Poluektov, E. J. J. Groenen, G. W. Canters, H. Nar, A. Messerschmidt, *J. Am. Chem. Soc.* **1997**, *119*, 4726–4731.

[61] W. E. Antholine, P. M. Hanna, D. R. McMillin, *Biophys. J.* **1993**, *64*, 267–272.

[62] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42–55.

[63] M. D. Rabenstein, Y. K. Shin, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8239–8243.

[64] E. J. Hustedt, A. H. Beth, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 129–153.

[65] E. J. Hustedt, R. A. Stein, L. Sethaphong, S. Brandon, Z. Zhou, S. C. DeSensi, *Biophys. J.* **2006**, *90*, 340–56.

[66] P. P. Borbat, J. H. Freed, *Struct. Bond.* **2013**, *152*, 1–82.

[67] M. J. N. Junk, Electron Paramagnetic Resonance Theory, University of California, **2012**.

[68] S. Saxena, J. H. Freed, *Chem. Phys. Lett.* **1996**, *251*, 102–110.

[69] G. Jeschke, M. Pannier, A. Godt, H. W. Spiess, *Chem. Phys. Lett.* **2000**, *331*, 243–252.

[70] V. V. Kurshev, A. M. Raitsimring, Y. D. Tsvetkov, *J. Magn. Reson.* **1989**, *81*, 441–454.

[71] A. Raitsimring, J. Peisach, H. C. Lee, X. Chen, *J. Phys. Chem.* **1992**, *96*, 3526–3531.

[72] G. M. Zhidomirov, K. M. Salikhov, *Sov. Phys. JETP* **1969**, *29*, 1037–1040.

[73] A. V. Kulikov, G. I. Likhtenstein, *Adv. Mol. Relax. Interact. Process.* **1977**, *10*, 47–79.

[74] M. H. Rakowsky, K. M. More, A. V Kulikov, G. R. Eaton, S. S. Eaton, *J. Am. Chem. Soc.* **1995**, *117*, 2049–2057.

[75] R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H. W. Spiess, *Angew. Chemie Int. Ed.* **1998**, *37*, 2833–2837.

[76] E. R. Georgieva, P. P. Borbat, C. Ginter, J. H. Freed, O. Boudker, O. Georgieva, E. R., Borbat, P. P., Ginter, C., Freed, J. H., Boudker, *Nat. Struct. Mol. Biol.* **2013**, *20*, 215–221.

[77] G. Jeschke, *ChemPhysChem* **2002**, *3*, 927–932.

[78] D. E. Kaplan, E. L. Hahn, *J. Phys. le Radium* **1958**, *19*, 821–825.

[79] P. P. Borbat, E. R. Georgieva, J. H. Freed, *J. Phys. Chem. Lett.* **2013**, *4*, 170–175.

[80] C. E. Tait, S. Stoll, *Phys. Chem. Chem. Phys.* **2016**, *18*, 18470–18485.

[81] C. L. Motion, J. E. Lovett, S. Bell, S. L. Cassidy, P. A. S. Cruickshank, D. R. Bolton, R. I. Hunter, H. El Mkami, S. Van Doorslaer, G. M. Smith, *J. Phys. Chem. Lett.* **2016**, *7*, 1411–1415.

[82] P. E. Spindler, S. J. Glaser, T. E. Skinner, T. F. Prisner, *Angew. Chemie Int. Ed.* **2013**, *52*, 3425–3429.

[83] G. Bodenhausen, R. L. Vold, R. R. Vold, *J. Magn. Reson.* **1980**, *37*, 93–106.

[84] J. G. Powles, P. Mansfield, *Phys. Lett.* **1962**, *2*, 58–59.

[85] P. Schöps, P. E. Spindler, A. Marko, T. F. Prisner, *J. Magn. Reson.* **2015**, *250*, 55–62.

[86] L. V. Kulik, Y. A. Grishin, S. A. Dzuba, I. A. Grigoryev, S. V. Klyatskaya, S. F. Vasilevsky, Y. D. Tsvetkov, *J. Magn. Reson.* **2002**, *157*, 61–68.

[87] S. Milikisyants, F. Scarpelli, M. G. Finiguerra, M. Ubbink, M. Huber, *J. Magn. Reson.* **2009**, *201*, 48–56.

[88] Y.-W. Chiang, P. P. Borbat, J. H. Freed, *J. Magn. Reson.* **2005**, *172*, 279–295.

[89] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473–498.

[90] A. Godt, M. Schulte, H. Zimmermann, G. Jeschke, *Angew. Chemie Int. Ed.* **2006**, *45*, 7560–7564.

[91] D. Margraf, B. E. Bode, A. Marko, O. Schiemann, T. F. Prisner, *Mol. Phys.* **2007**, *105*, 2153–2160.

[92] D. Margraf, P. Cekan, T. F. Prisner, S. T. Sigurdsson, O. Schiemann, *Phys. Chem. Chem. Phys.* **2009**, *11*, 6708–6714.

[93] A. Savitsky, A. A. Dubinskii, M. Flores, W. Lubitz, K. Möbius, *J. Phys. Chem. B* **2007**, *111*, 6245–6262.

[94] G. W. Reginsson, O. Schiemann, *Biochem. J.* **2011**, *434*, 353–363.

[95] G. W. Reginsson, R. I. Hunter, P. A. S. Cruickshank, D. R. Bolton, S. T. Sigurdsson, G. M. Smith, O. Schiemann, *J. Magn. Reson.* **2012**, *216*, 175–182.

[96] V. P. Denysenkov, T. F. Prisner, J. Stubbe, M. Bennati, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13386–13390.

[97] V. P. Denysenkov, T. F. Prisner, J. Stubbe, M. Bennati, D. Biglino, W. Lubitz, T. F. Prisner, M. Bennati, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1224–1227.

[98] B. E. Bode, J. Plackmeyer, T. F. Prisner, O. Schiemann, *J. Phys. Chem. A* **2008**, *112*, 5064–5073.

[99] J. E. Lovett, A. M. Bowen, C. R. Timmel, M. W. Jones, J. R. Dilworth, D. Caprotti, S. G. Bell, L. L. Wong, J. Harmer, *Phys. Chem. Chem. Phys.* **2009**, *11*, 6840–8.

[100] A. M. Bowen, M. W. Jones, J. E. Lovett, T. G. Gaule, M. J. McPherson, J. R. Dilworth, C. R. Timmel, J. R. Harmer, *Phys. Chem. Chem. Phys.* **2016**, *18*, 5981–5994.

[101] C. Abé, D. Klose, F. Dietrich, W. H. Ziegler, Y. Polyhach, G. Jeschke, H.-J. Steinhoff, *J. Magn. Reson.* **2012**, *216*, 53–61.

[102] I. Tkach, S. Pornsuwan, C. Höbartner, F. Wachowius, S. T. Sigurdsson, T. Y. Baranova, U. Diederichsen, G. Sicoli, M. Bennati, *Phys. Chem. Chem. Phys.* **2013**, *15*, 3433–3437.

[103] A. Marko, T. F. Prisner, *Phys. Chem. Chem. Phys.* **2013**, *15*, 619–627.

[104] S. Valera, B. E. Bode, *Molecules* **2014**, *19*, 20227–20256.

[105] A. D. Milov, Y. D. Tsvetkov, *Appl. Magn. Reson.* **1997**, *12*, 495–504.

[106] J. P. Klare, H.-J. Steinhoff, *Struct. Bond.* **2013**, *152*, 205–248.

[107] R. Ward, O. Schiemann, *Struct. Bond.* **2012**, 249–281.

[108] I. Krstić, R. Hänsel, O. Romainczyk, J. W. Engels, V. Dötsch, T. F. Prisner, *Angew. Chemie Int. Ed.* **2011**, *50*, 5070–5074.

[109] M. Qi, A. Groß, G. Jeschke, A. Godt, M. Drescher, *J. Am. Chem. Soc.* **2014**, *136*, 15366–15378.

[110] E. Narr, A. Godt, G. Jeschke, *Angew. Chemie Int. Ed.* **2002**, *41*, 3907–3910.

[111] B. E. Bode, J. Plackmeyer, M. Bolte, T. F. Prisner, O. Schiemann, *J. Organomet. Chem.* **2009**, *694*, 1172–1179.

[112] Z. Yang, D. Kise, S. Saxena, *J. Phys. Chem. B* **2010**, *114*, 6165–6174.

[113] M. Ezhevskaya, E. Bordignon, Y. Polyhach, L. Moens, S. Dewilde, G. Jeschke, S. Van Doorslaer, *Mol. Phys.* **2013**, *111*, 2855–2864.

[114] D. Akhmetzyanov, J. Plackmeyer, B. Endeward, V. P. Denysenkov, T. F. Prisner, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6760–6766.

[115] A. Meyer, O. Schiemann, *J. Phys. Chem. A* **2016**, *120*, 3463–3472.

[116] I. Kaminker, M. Bye, N. Mendelman, K. Gislason, S. T. Sigurdsson, D. Goldfarb, *Phys. Chem. Chem. Phys.* **2015**, *17*, 15098–102.

[117] P. Lueders, G. Jeschke, M. Yulikov, *J. Phys. Chem. Lett.* **2011**, *2*, 604–609.

[118] M. Yulikov, P. Lueders, M. Farooq Warsi, V. Chechik, G. Jeschke, *Phys. Chem. Chem. Phys.* **2012**, *14*, 10732.

[119] I. Kaminker, I. Tkach, N. Manukovsky, T. Huber, H. Yagi, G. Otting, M. Bennati, D. Goldfarb, *J. Magn. Reson.* **2013**, *227*, 66–71.

[120] L. Garbuio, E. Bordignon, E. K. Brooks, W. L. Hubbell, G. Jeschke, M. Yulikov, *J. Phys. Chem. B* **2013**, *117*, 3145–3153.

[121] S. Razzaghi, M. Qi, A. I. Nalepa, A. Godt, G. Jeschke, M. Yulikov, *J. Phys. Chem. Lett.* **2014**, 1–3.

[122] A. V. Bogachev, D. A. Bloch, Y. V. Bertsova, M. I. Verkhovsky, *Biochemistry* **2009**, 6299–6304.

[123] A. V. Astashkin, B. O. Elmore, W. Fan, J. G. Guillemette, C. Feng, *J. Am. Chem. Soc.* **2010**, *132*, 12059–12067.

[124] A. V. Astashkin, *Methods Enzymol.* **2015**, *563*, 251–284.

[125] A. V. Astashkin, J. Seravalli, S. O. Mansoorabadi, G. H. Reed, S. W. Ragsdale, *J. Am. Chem. Soc.* **2006**, *128*, 3888–3889.

[126] A. V. Astashkin, A. Rajapakshe, M. J. Cornelison, K. Johnson-Winters, J. H. Enemark, *J. Phys. Chem. B* **2012**, *116*, 1942–1950.

[127] A. V. Astashkin, L. Chen, X. Zhou, H. Li, T. L. Poulos, K. J. Liu, J. G. Guillemette, C. Feng, *J. Phys. Chem. A* **2014**, *118*, 6864–6872.

[128] A. Savitsky, J. Niklas, J. H. Golbeck, K. Möbius, W. Lubitz, *J. Phys. Chem. B* **2013**, *117*, 11184–11199.

[129] A. Savitsky, A. A. Dubinskii, H. Zimmermann, W. Lubitz, K. Möbius, *J. Phys. Chem. B* **2011**, *115*, 11950–11963.

[130] S. Razzaghi, E. K. Brooks, E. Bordignon, W. L. Hubbell, M. Yulikov, G. Jeschke, *ChemBioChem* **2013**, *14*, 1883–1890.

[131] A. Collauto, V. Frydman, M. D. Lee, E. H. Abdelkader, A. Feintuch, J. D. Swarbrick, B. Graham, G. Otting, D. Goldfarb, *Phys. Chem. Chem. Phys.* **2016**, *18*, 19037–19049.

[132] B. Svensson, T. Oda, F. R. Nitu, Y. Yang, I. Cornea, Y. Chen-Izu, J. D. Fessenden, D. M. Bers, D. D. Thomas, R. L. Cornea, *Biophys. J.* **2014**, *107*, 2037–2048.

[133] H. C. Hyde, W. Sandtner, E. Vargas, A. T. Dagcan, J. L. Robertson, B. Roux, A. M. Correa, F. Bezanilla, *Structure* **2012**, *20*, 1629–1640.

[134] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, **1992**.

[135] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics* **2006**, *22*, 195–201.

[136] X. J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108–5121.

[137] K. Sale, L. Song, Y. S. Liu, E. Perozo, P. Fajer, *J. Am. Chem. Soc.* **2005**, *127*, 9334–9335.

[138] Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356–2366.

[139] G. Hagelueken, R. Ward, J. H. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377–391.

[140] M. M. Hatmal, Y. Li, B. G. Hegde, P. B. Hegde, C. C. Jao, R. Langen, I. S. Haworth, *Biopolymers* **2012**, *97*, 35–44.

[141] E. G. B. Evans, M. J. Pushie, K. A. Markham, H.-W. Lee, G. L. Millhauser, *Structure* **2016**, *24*, 1057–1067.

[142] D. M. J. Liley, F. Eckstein, *Ribozymes and RNA Catalysis*, The Royal Society Of Chemistry, Cambridge, **2007**.

Appendix

[P1] Geometric model based fitting algorithm for orientation selective PELDOR data

Reprinted with permission from

Dinar Abdullin, Gregor Hagelueken, Robert I. Hunter, Graham M. Smith, Olav Schiemann, "Geometric Model Based Fitting Algorithm for Orientation Selective PELDOR Data", *Mol. Phys.* **2015**, *113*, 544-560. DOI: 10.1080/00268976.2014.960494

Own contribution to the manuscript:

- Development and implementation of the algorithm
- Application of the algorithm to test data
- Writing the manuscript

Molecular Physics, 2015
Vol. 113, No. 6, 544–560, http://dx.doi.org/10.1080/00268976.2014.960494

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Geometric model-based fitting algorithm for orientation-selective PELDOR data

Dinar Abdullin[a], Gregor Hagelueken[a], Robert I. Hunter[b], Graham M. Smith[b] and Olav Schiemann[a,*]

[a]*Institute of Physical and Theoretical Chemistry, University of Bonn, Bonn, Germany;* [b]*School of Physics, University of St Andrews, St Andrews, UK*

(*Received 21 March 2014; accepted 27 August 2014*)

Pulsed electron–electron double resonance (PELDOR or DEER) spectroscopy is frequently used to determine distances between spin centres in biomacromolecular systems. Experiments where mutual orientations of the spin pair are selectively excited provide the so-called orientation-selective PELDOR data. This data is characterised by the orientation dependence of the modulation depth parameter and of the dipolar frequencies. This dependence has to be taken into account in the data analysis in order to extract distance distributions accurately from the experimental time traces. In this work, a fitting algorithm for such data analysis is discussed. The approach is tested on PELDOR data-sets from the literature and is compared with the previous results.

Keywords: DEER; spin labels; metal centres; simulations

1. Introduction

The measurement of inter-spin distances on the nanometre scale by pulsed electron–electron double resonance (PELDOR or DEER) [1,2] is a powerful tool to gain structural information about proteins and oligonucleotides [3–7]. The increasing variety of molecular systems and types of spin centres studied by PELDOR requires further development of the method. One of the challenges for PELDOR is related to an accurate extraction of distance distributions from the experimental time-domain data. Two scenarios for the data analysis are usually differentiated. In the first case, it is assumed that all orientations of the inter-spin distance vector $\boldsymbol{r}$ are contributing to the PELDOR time trace with weighting that follows the powder distribution. If this condition is fulfilled, the Fourier transformation of the experimental signal yields a dipolar spectrum referred to as the Pake pattern, and the distance distribution can be reconstructed from the experimental PELDOR time trace by using Tikhonov regularisation as implemented in the DeerAnalysis program [8]. This analysis covers the majority of PELDOR experiments performed on nitroxide spin labels due to the inherent flexibility of the linker connecting the labels with, e.g., a protein. The second case corresponds to those time traces for which only certain orientations of the spin–spin vector $\boldsymbol{r}$ relative to the external magnetic field $\boldsymbol{B_0}$ were excited. Examples for such a case are the tyrosine radicals in the ribonucleotide reductase dimer [9,10], the rigid spin label Ç in oligonucleotides [11,12], spin labels restricted in their rotational freedom by the protein environment [13,14] or transition metal ions/clusters [15–22]. In each case, the angle θ between the inter-spin vector $\boldsymbol{r}$ and the external magnetic field $\boldsymbol{B_0}$ is randomly distributed within the sample, but the orientation selectivity is obtained because: (1) the mutual orientation of both spin centres is nearly the same in each molecule due to the inhibited motional freedom of the spin centres and (2) the pulse sequence excites only certain orientations of the spin pairs due to the narrow excitation profile of the pulses with respect to the width of the spectrum. In contrast, a bisnitroxide embedded into a crystal lattice can show only one θ angle depending on the orientation of the crystal in the magnetic field even if the whole spectrum were to be excited [23]. In all of the examples above, the dipolar spectrum deviates from the Pake pattern and therefore, alternative methods are required for extracting the distance distributions.

In recent years, several approaches were proposed for the analysis of orientation selective PELDOR data. In most of these works, PELDOR time traces were collected at several magnetic fields and/or frequency offsets in order to probe different orientations of spin centres. If the majority of the orientations are probed, summing up all time traces can reduce the effect of orientation selection. In this case, the easiest way to obtain the inter-spin distance is to analyse the resulting averaged time trace by Tikhonov regularisation [24]. This approach usually provides a good estimate of the inter-spin distance. However, the width of the distance distribution cannot be determined with high accuracy and all information about the orientation of spin centres is lost.

* Corresponding author. Email: schiemann@pc.uni-bonn.de

In a more elaborate approach, a molecular model defining the orientations of the spin centres is constructed and used to simulate the PELDOR time traces. Several models were used to simulate PELDOR data of nitroxide biradicals [25–30], amino acid biradicals [9,10], trityl-nitroxide biradicals [31], and metal-nitroxide as well as metal–metal complexes [17,21]. The initial geometry of the model was usually based on structural information obtained from X-ray diffraction measurements and/or density functional theory (DFT) and molecular dynamics (MD) calculations. Note that the availability of such information is crucial for the method. Then, the geometric parameters of the model were varied until the simulated PELDOR time traces matched the experimental ones, judged by eye. As a result, this method provides the relative orientation of spin centres and, in some cases, the flexibility of the molecular linker connecting them in addition to the distances. A drawback of this method is that a large space of parameters has to be scanned until the optimal geometry of the model is found. Unless an optimisation algorithm is applied for this, this task becomes very time consuming and it is not clear whether the solution obtained is unique. In the papers of Abé *et al*. [14] and Tkach *et al*. [32], grid search and sequential quadratic programming optimisation algorithms, respectively, were used for rigid spin systems. For these systems, the number of parameters could be reduced down to three [14] and six [32], respectively. This allowed exploration of the whole possible range of each optimisation parameter within a reasonable time. Moreover, these studies revealed that several sets of optimisation parameters deliver very similar fits to the experimental PELDOR data due to the symmetry of the problem.

A model-free approach was proposed by Marko *et al*. [33]. Their fitting algorithm reconstructs an experimental data-set by searching for an optimal combination of pre-simulated PELDOR time traces of all relative orientations and inter-spin distances in the experimentally accessible range. Good fits to experimental PELDOR data were obtained on the model bisnitroxide systems investigated. The authors pointed out that the solutions were not unique, which might be a consequence of the insufficient orientation selectivity achieved at the X-band frequencies.

In this study, a different fitting algorithm for the analysis of orientation-selective PELDOR data is presented. It aims to extend the application of the model-based data analysis for systems with unknown structures and dynamics. To do this, a simplified model of a spin pair is created and optimised in order to provide the best fit to PELDOR time traces. The optimisation is performed by a genetic algorithm. This algorithm has shown its ability to explore a large parameter space and to find the global minimum of a problem in a wide range of applications, including electron paramagnetic resonance (EPR) data analysis [34–36]. The algorithm was tested on the PELDOR data-sets stemming from nitroxide biradicals **1** and **2** measured at W-band [30] and from trityl-nitroxide biradical **3** measured at X-band [31] (Figure 1). The geometric models obtained here are then compared to the models from the original papers.

2. Materials and methods

The synthesis of nitroxide biradicals **1** and **2** and the corresponding PELDOR measurements have been described elsewhere [27,30,37]. The time traces of **1** were acquired at W-band frequencies using a unique home-built spectrometer that features 1 kW output powers and uses a non-resonant sample-holder design that allows the use of large volume samples (similar volumes to X-band) [38]. Despite being non-resonant, it features $\pi/2$ pulse lengths comparable to X-band and offers a large increase in concentration sensitivity. This instrument has a 1 GHz instantaneous bandwidth that allows pump and probe pulses to be positioned anywhere across the W-band nitroxide spectrum, where the *g*-tensor is now fully resolved. In order to resolve the orientation selection, six different combinations of the pump and detection frequencies were applied. Each frequency was set to be in resonance with either X, Y, or Z spectral component of the nitroxide *g*-tensor, resulting in six possible combinations for the nitroxide spin pair: XX, YY, ZZ, YX, YZ, and ZX. The exact magnetic field positions, pump and probe frequencies, and other experimental parameters are given in Ref. [30]; the ones required here are collected in the Supporting Material.

The synthesis of trityl-nitroxide biradical **3** and the PELDOR measurements performed on this molecule are described in Ref. [31]. In the PELDOR measurements, the pump pulse was applied on the trityl spin, whereas the echo signal was detected on the different spectral positions of the nitroxide. In total, seven different positions in the nitroxide spectrum were probed resulting in the frequency offsets from 30 to 90 MHz. Note that due to the low anisotropy of the trityl spectrum at X-band, the 16 ns pump pulse used in the experiments allowed a complete inversion of the trityl spin resulting in a very deep modulation of the PELDOR time traces. For more experimental details, see Ref. [31] and Section 1 of the Supporting Material.

The background correction and the Tikhonov regularisation of the PELDOR time traces were done by means of the DeerAnalysis2013 program [8]. Data files with the background corrected time traces were then used as input data in the following analysis.

The fitting algorithm for the orientation-selective PELDOR data analysis was implemented in C++ in the program PeldorFit, which is available at http://www.schiemann.uni-bonn.de. The performance of the program was tested on a 64-core workstation from sysGen GmbH with 2.3 GHz processor frequency and 132 GB RAM.

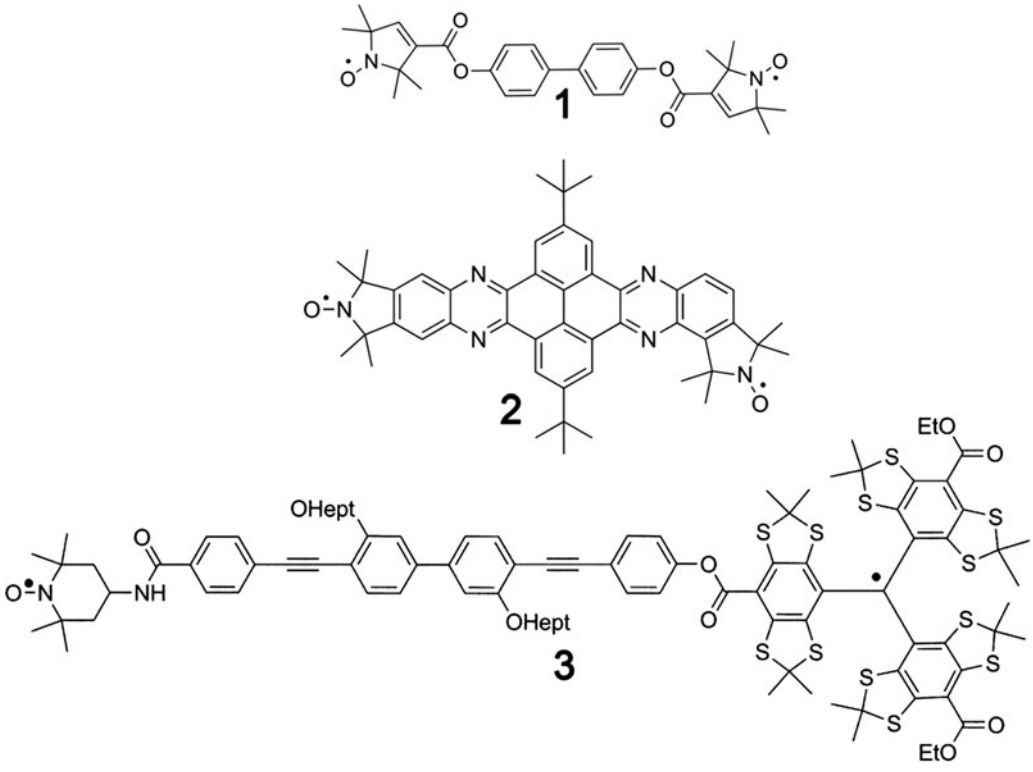


Figure 1. Structures of the nitroxide biradicals **1** and **2** and trityl-nitroxide biradical **3**.

3. Theory

3.1. PELDOR methodology

The scheme of the 4-pulse sequence of the PELDOR experiment is shown in Figure 2. The pulses at frequency ν_{det} (detection frequency) create a refocused spin echo from those spins that are in resonance with these pulses; these spins are called A spins. Introduction of an inversion pulse with frequency ν_{pump} (pump frequency) flips the spins in resonance with this frequency; these spins are called B spins. Due to the dipole–dipole interaction between A and B spins, the flip of the B spins changes the local magnetic field at the A spins. As a consequence, the Larmor frequency of the A spins is shifted by

$$\omega_{dd} = \frac{\mu_0 \beta_e^2 g_A g_B}{4\pi\hbar} \frac{(1 - 3\cos^2\theta)}{\boldsymbol{r}^3}. \quad (1)$$

Here, μ_0 is the vacuum permeability, β_e is the Bohr magneton, $\hbar$ is the reduced Planck constant, g_A and g_B are the g-factors of spins A and B, respectively, $\boldsymbol{r}$ is the length of the radius vector connecting the two spins A and B, and θ is the angle between $\boldsymbol{r}$ and the external magnetic field B_0. Depending on the sign of the frequency shift, the A spins accumulate a phase shift of $\omega_{dd}t$ or $-\omega_{dd}t$. Here, t defines the time position of the pump pulse with respect to the primary echo. The total magnetisation of these two sub-ensembles oscillates as $\exp(i\omega_{dd}t) + \exp(-i\omega_{dd}t) \sim \cos(\omega_{dd}t)$. Thus, for different t values, the primary echo is refocused with time-dependent phases and is modulated by the dipolar frequencies.

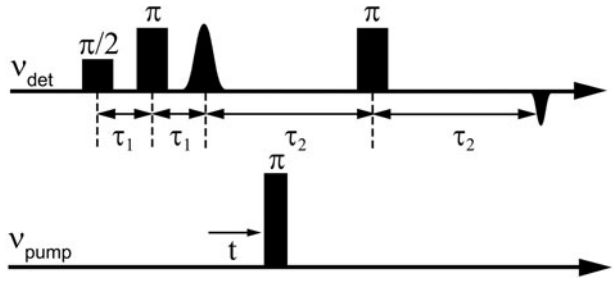


Figure 2. The pulse sequence of the PELDOR experiment.

The time dependence of the echo intensity in the PELDOR experiment can be described by a product of two contributions:

$$V(t) = V_{intra}(t) \cdot V_{inter}(t). \quad (2)$$

$V_{intra}(t)$ is the contribution of the dipolar interaction between spin centres within the same molecule or molecular assembly, whereas $V_{inter}(t)$ denotes the echo decay caused

by the intermolecular distribution of the spin centres in the sample. In the case of a homogenous distribution, $V_{\text{inter}}(t)$ is a monoexponential decay and is removed during the data analysis. The intramolecular contribution is given by

$$V_{\text{intra}}(t) = \prod_{i=1}^{N_c} V_i(0)\left(1 - \int_0^{\pi/2} \lambda_i(\theta)\,[1 - \cos(\omega_{\text{dd}}(r_i, \theta)t)]\,\sin(\theta)\text{d}\theta\right), \quad (3)$$

where the summation is done over all N_c conformers of the spin pair. For the ith conformer, r_i is the distance between the spins A and B, $V_i(0)$ is the spin A echo intensity at $t = 0$, $\lambda_i(\theta)$ is the probability of the spin B being flipped by the pump pulse and is called the modulation depth parameter. The whole ensemble of the inter-spin distances r_i determine the distance distribution function $D(r)$, which is the goal of the PELDOR data analysis. If the correlation between the orientation of the spin centres and the inter-spin distance can be neglected, Equation (3) can be transformed into [39]

$$V_{\text{intra}}(t) = V(0)\left(1 - \int_0^{\infty} D(r) \int_0^{\pi/2} \lambda(\theta)\,[1 - \cos(\omega_{\text{dd}}(r, \theta)t)] \cdot \sin(\theta)\text{d}\theta\text{d}r\right), \quad (4)$$

where $V(0)$ and $\lambda(\theta)$ are the spin A echo intensity and the modulation depth parameter of the whole system of N_c conformers, respectively.

In orientationally disordered systems, when the correlation between the orientations of spin centres can be neglected, the modulation depth parameter λ is independent of θ. Under this assumption, the distance distribution can be deduced from the time-domain signal $V_{\text{intra}}(t)$ via Tikhonov regularisation as implemented in DeerAnalysis. However, if the angular correlation between spin centres is significant, this simplification is not valid. Thus, in order to extract the accurate distance distribution from the PELDOR data, the $\lambda(\theta)$ dependence has to be included in the analysis. This case is referred to in the literature as orientation-selective PELDOR. Its intricacy arises from the fact that $\lambda(\theta)$ is usually unknown.

3.2. *Orientation-selective PELDOR data analysis*

In this data analysis, a simplified geometric model of the spin system is created and optimised in order to provide the best fit to the experimental PELDOR time traces. Later on, the unknown $D(r)$ and $\lambda(\theta)$ distributions are derived from the optimised model. In the following, the geometric model of the spin system, the calculation of the PELDOR time traces for this model, and the fitting of the experimental PELDOR data are successively outlined.

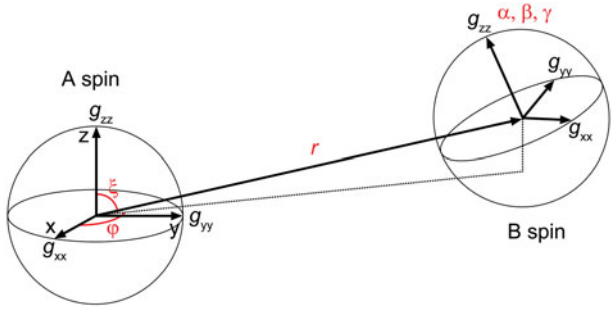


Figure 3. Geometric model of the spin pair. The g-tensor of the A spin is supposed to coincide with the coordinate system of the model. The g-tensor of the B spin is defined by the Euler angles (α, β, γ) relative to spin A. The Euler angles are defined in accordance with the z–x'–z'' convention. The inter-spin vector connecting the two spin centres is represented by the spherical coordinates (r, ξ, φ) relative to spin A. All geometric parameters have a distribution. For simplicity, the $D(r)$ distribution is approximated by the normal distribution; α, β, γ, ξ, and φ angles are supposed to be equally distributed in the intervals $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$, $\Delta\xi$, and $\Delta\varphi$, respectively.

The geometric model used for the fitting is shown in Figure 3. It consists of two coordinate frames associated with the g-tensor principle axes of the two spins forming the spin pair. Both spins are considered as point objects, i.e. spin delocalisation is neglected. For clarity, the naming of the spins is chosen in accordance with the PELDOR theory (see above). The reference coordinate system of the model is set to be coincident with the g-tensor principle axes of the spin A. Then, the orientation of the g-tensor principle axes corresponding to the spin B is represented by three Euler angles (α, β, γ). The Euler angles are defined in accordance with the z–x'–z'' convention. The vector connecting the origins of these two frames determines the inter-spin vector $\boldsymbol{r}$. In the reference coordinate system, $\boldsymbol{r}$ is described by spherical coordinates: length of the vector $\boldsymbol{r}$, polar ξ, and azimuthal φ angles. Due to the intrinsic flexibility of each molecule, the described geometric parameters usually have some distribution. The distribution of the inter-spin distances is approximated by the normal distribution, and the α, β, γ, ξ, and φ angles are assumed to be uniformly distributed in the intervals $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$, $\Delta\xi$, and $\Delta\varphi$, respectively. As a result, in total, 12 parameters are used to describe the model: the mean inter-spin distance μ, its standard deviation σ, and the mean value and distribution width of the five angles. Note that all parameters of the model are assumed to be independent of each other. This condition differs from the studies, where more specific modes of motion were included with correlations between the distance and relative orientation of the spin centres as well as correlations between the different Euler angles [17,30]. Such correlations are neglected here, because the structure is

Table 1. Number and ranges of Euler angles (α, β, γ) and polar angles (ξ, φ) required to simulate the PELDOR time traces in the cases of isotropic, axial, and rhombic magnetic tensors of spins A and B.

Spin A	Spin B	α	β	γ	ξ	φ
Axial	Isotropic	–	–	–	[0°, 90°]	–
Axial	Axial	[0°, 180°]	[0°, 90°]	–	[0°, 90°]	–
Rhombic	Isotropic	–	–	–	[0°, 90°]	[0°, 180°]
Rhombic	Axial	[0°, 180°]	[0°, 90°]	–	[0°, 90°]	[0°, 180°]
Rhombic	Rhombic	[0°, 180°]	[0°, 90°]	[0°, 180°]	[0°, 90°]	[0°, 180°]

unknown. However, the distributions of the angular parameters are taken into consideration here, which differs from the previous works where a similar model was established for rigid spin labels [14,32].

Depending on the symmetry of the g- and A-tensors (magnetic tensors) of spins A and B, all 12 geometric parameters as defined above or only a sub-set of them are needed to simulate the PELDOR time trace. If the magnetic tensors of both the spins are rhombic, the full set of parameters is required for the calculations. In contrast, eight parameters are sufficient when the magnetic tensors of both the spins are axial. For example, at X-band frequencies, the g- and A-tensors of the nitroxides can be considered as axial, which allows one to exclude the azimuthal angle φ and one of the Euler angles. Moreover, due to the inversion symmetry of the magnetic tensors (spin Hamiltonian), there are a number of relative orientations of the spin centres, which provide identical PELDOR time traces. If each spin is considered as a point object, the 180° rotation of one of the g-tensor principle axes does not influence the shape of the PELDOR time trace [15]. All these considerations are summarised in Table 1, which specifies the number of required angular parameters and their variation ranges for all possible combinations of isotropic, axial, and rhombic spin systems A and B.

After the set of geometric parameters has been chosen, the PELDOR time trace can be calculated according to Equation (4). The double integral in Equation (4) is evaluated via the Monte-Carlo (MC) method, setting the number of samples to 10^5. The $D(r)$ function is approximated by the normal distribution with the mean value μ and standard deviation σ. The calculation of the $\lambda(\theta)$ dependence is made in accordance with the previous works [30,40] and includes the following steps. First, the θ angle between the external magnetic field $\boldsymbol{B}_0$ and the inter-spin vector $\boldsymbol{r}$ has to be determined in each MC step. For this purpose, one conformer of the spin pair is created by randomly choosing values of r, ξ, φ, α, β, and γ from their distributions. Then, the unit vector $\boldsymbol{B}_0/B_0$ of the magnetic field is introduced in the coordinate system of the model. This vector has a random orientation with respect to the inter-spin vector $\boldsymbol{r}$. As soon as the $\boldsymbol{B}_0$ and $\boldsymbol{r}$ vectors are determined within the same coordinate frame, the θ angle is found from their mutual orientation.

Second, the resonance frequencies of the A and B spins are calculated for the generated conformer of the spin pair and magnetic field vector $\boldsymbol{B}_0$. It is assumed that the resonance frequencies of both the spins are determined by the Zeeman interaction, the hyperfine interaction of the electron spins, and inhomogeneous line broadening. These contributions are usually dominant in the spin Hamiltonian of the majority of paramagnetic centres. Unresolved anisotropy of the g- and A-tensors can also be taken into account via g-strain and A-strain as it is done in the EasySpin program [36].

Third, the transverse magnetisation of spin A and the inversion probability of spin B are calculated for each generated conformer of the spin pair. These probabilities depend on the proximity of the resonance frequencies of the A and B spins to the frequencies of the microwave pulses. According to Ref. [40], they can be calculated as

$$p_{\mathrm{det}}(\omega) = \frac{\omega_1'}{\Omega'} \sin\left(\Omega' t_{\pi/2}\right) \frac{\omega_1^4}{4\Omega^4} \left[1 - \cos\left(\Omega t_\pi\right)\right]^2,$$
$$p_{\mathrm{pump}}(\omega) = \frac{\omega_1^2}{2\Omega^2} \left[1 - \cos\left(\Omega t_\pi\right)\right], \tag{5}$$

where

$$\Omega^2 = \omega_1^2 + (\omega - \omega_{\mathrm{mw}})^2,$$
$$\Omega'^2 = \omega_1'^2 + (\omega - \omega_{\mathrm{mw}})^2,$$
$$\omega_1 = \frac{\pi}{t_\pi},$$
$$\omega_1' = \frac{\pi/2}{t_{\pi/2}}. \tag{6}$$

Here, ω is the resonance frequency of the A or B spin, $t_{\pi/2}$ and t_π are the lengths of the $\pi/2$ and π pulses, respectively. The probability p_{det} is calculated for the lengths of the detection pulses and $\omega_{\mathrm{mw}} = 2\pi\nu_{\mathrm{det}}$, whereas the length of the inversion pulse and $\omega_{\mathrm{mw}} = 2\pi\nu_{\mathrm{pump}}$ are used to compute p_{pump}. In this study, p_{pump} is multiplied by a constant factor, $0 < \eta \leq 1$, which takes into account factors affecting the modulation depth parameter: non-ideal pulses and incomplete spin labelling, etc. The calculated probabilities are then used to calculate the modulation depth parameter for

the individual *i*th conformer of the spin pair:

$$\lambda^i = \frac{p^i_{\mathrm{det}}(\omega_A)\cdot p^i_{\mathrm{pump}}(\omega_\mathrm{B}) + p^i_{\mathrm{det}}(\omega_B)\cdot p^i_{\mathrm{pump}}(\omega_A)}{\sum_{j=1}^{N_\mathrm{MC}}\left(p^j_{\mathrm{det}}(\omega_A) + p^j_{\mathrm{det}}(\omega_B)\right)}, \quad (7)$$

where ω_A and ω_B are the resonance frequencies of the A and B spins, respectively, and N_MC is the number of generated conformers, which is equal to the number of MC samples. Combinations of the calculated θ and λ values for the whole set of conformers determine the $\lambda(\theta)$ dependence. As a result, all components of the integrand in Equation (4) have been determined, which allows one to calculate the PELDOR time trace for the constructed model.

Now, the deviation of the calculated time trace from the experimental one is minimised using the genetic algorithm [41]. This algorithm can be used to solve optimisation problems. Here, the geometric parameters of the model are varied until the corresponding calculated PELDOR time traces provide the best fit to the experimental ones. In this problem, the geometric parameters play the role of fitting parameters. In addition, one has to decide which characteristic will be used to judge the goodness of the fit. Probably, the most commonly used one is the root-mean square deviation (RMSD):

$$\mathrm{RMSD} = \sqrt{\frac{1}{N_t}\sum_{t=1}^{N_t}(V_\mathrm{exp}(t) - V_\mathrm{calc}(t))^2}. \quad (8)$$

Here, $V_\mathrm{exp}(t)$ is an experimental background corrected time trace and $V_\mathrm{calc}(t)$ is its calculated prediction. N_t is the number of the data points in the experimental data-set.

The genetic algorithm was chosen because: (1) it is a global optimisation algorithm, i.e. it is able to find a global minimum of the problem even when the problem has several local minima, (2) it does not require initial guesses and any derivative information, and (3) it is well suited to deal with a large number of fitting parameters in terms of computational time [41]. In the following, a brief description of the implemented algorithm is given.

In the terminology of the genetic algorithm, each fitting parameter is called a 'gene'. A set of genes determines one complete set of fitting parameters and is named a 'chromosome'. An evolutionary cycle begins with the creation of M chromosomes forming a first 'generation'. The genes of the initial chromosomes are generated randomly in certain ranges. For example, the gene responsible for the mean distance μ can be chosen in the range from 2 to 4 nm, whereas the ξ angle is varied between 0° and 90°. After a set of chromosomes is created, each of the chromosomes is used to calculate the PELDOR signal, according to Equation (4) as described above. Subsequently, the RMSD is calculated for each chromosome. In terms of the genetic algorithm, this calculation determines the scoring process. The scoring process assigns a goodness-of-fit value to each individual chromosome as a 'fitness' property. For example, a chromosome with higher fitness will have a lower RMSD value. Moreover, in accordance with the evolutionary principle, it determines which chromosomes will be used to produce new chromosomes. Thus, the next milestone of the evolutionary cycle is the production of an offspring. This operation involves three processes – selection of two chromosomes (parents) out of the set of M chromosomes, crossover of genes between them with probability p_c, and mutation of genes with probability p_m. The picking of the 'parents' is done by randomly choosing two pairs of chromosomes and selecting from each pair the chromosome of highest fitness (tournament selection). This ensures that the chromosomes with higher fitness will have higher chances of being chosen. This procedure is done $M/2$ times resulting in a new set of M modified chromosomes. This set of chromosomes is used as a starting point for the next evolutionary cycle, continued by scoring and mating as described above. This process is repeated until the desired number of generations, N_max is reached.

The number of chromosomes M, the maximal number of generations N_max, and the crossover and mutation probabilities p_c and p_m are the parameters to be set. These parameters may influence the ability of the genetic algorithm to find the global minimum of the problem. Therefore, the parameters of the genetic algorithm have to be adjusted. However, there is no strict rule of how to determine the optimal values of these parameters and usually they have to be found empirically. Moreover, the optimal values of the parameters may vary depending on the particular PELDOR data-set. Our tests showed that reasonable fits to the PELDOR data could be obtained with $M = 128$ (about 10 chromosomes per 1 optimisation parameter), $N_\mathrm{max} = 500$, $p_c = 0.3$, and $p_m = 0.05$. For the nitroxide biradical **2**, the number of chromosomes M was increased up to 192, since two additional fitting parameters were used.

Since all fitting parameters are only characteristics of the molecular model, the time traces recorded at various pump and detection frequencies can be fitted simultaneously (global fit). In this case, the goodness-of-fit parameters for the individual time traces are summed together in each optimisation step, and this sum is minimised.

4. Results and discussion

The fitting program was benchmarked on the PELDOR data-sets of biradicals **1**, **2**, and **3**. All data-sets consist of multiple time traces acquired at different frequency offsets $\Delta\nu$ and demonstrate pronounced orientation selectivity. Each time trace contains two constraints, dipolar frequencies and modulation depth, yielding a total of 12 constraints for the fitting algorithm for **1** and **2** and 14 for **3**. All time traces for each molecule were fitted together in a global fit in order to provide a geometric model, which is compatible with the whole data-set. The experimental PELDOR data were also analysed in previous works via different

molecular models and simulations. This allows making a qualitative comparison of the fitting results obtained here with the existing models.

4.1. Nitroxide biradical 1

The fitting results for bisnitroxide **1** are presented in Figure 4. As can be seen, good fits of the time traces were achieved. The frequency of the dipolar modulation, its damping, and the modulation depth are well reproduced for each experimental time trace. The corresponding optimised parameters of the geometric model are listed in Table 2. Among them, an inter-spin distance of 1.94 nm and a standard deviation of 0.03 nm were determined. This distance is in good agreement with the one found in the previous study ($\mu = 1.93$ nm and $\sigma = 0.03$ nm [30]). The distance distribution is also compared to the one obtained after the Tikhonov regularisation of the time trace

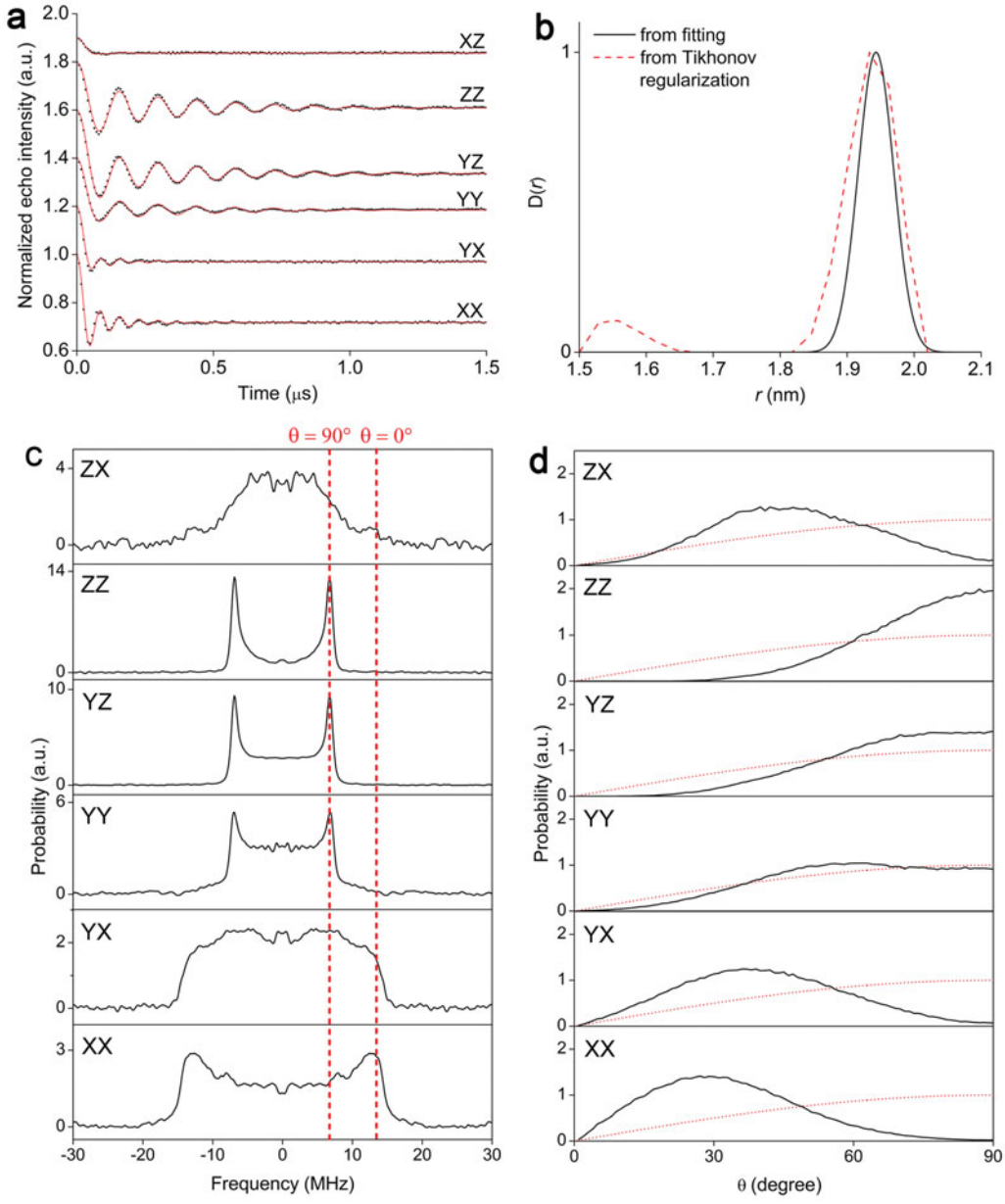


Figure 4. Fitting results for the PELDOR time traces of **1**. (a) The experimental time traces (dots) overlaid with their fits (lines). (b) The inter-spin distance distributions obtained from the fitting (solid line) and from the Tikhonov regularisation of the time trace averaged over all offsets (dash line). (c) The dipolar spectra obtained after Fourier transformation of each of the six experimental time traces. The positions of the parallel and perpendicular components are marked by red dashed lines. (d) The form factors $P(\theta)$ calculated for the optimised geometric model are shown for each time trace. A $\sin\theta$ distribution is depicted by a red dashed line and corresponds to the case without orientation selection.

Table 2. Optimized parameters of the geometric models of biradicals 1, 2, and 3.

biradical	**1**	**2**	**3**
μ, σ (nm)[a]	1.94, 0.03 (0.01, 0.01)	1.80, 0.02 (0.04, 0.04)	3.46, 0.06 (0.04, 0.04)
ξ, $\Delta\xi$ (°)[a]	76, 25 (21, 48)	79, 22 (17, 61)	74, 11 (25, 90)
φ, $\Delta\varphi$ (°)[a]	25, 49 (52, 81)	46, 3 (10, 27)	–
α, $\Delta\alpha$ (°)[a]	19, 3 (4, 39)	26, 48 (1, 1)	–
β, $\Delta\beta$ (°)[a]	52, 76 (5, 24)	3, 6 (7, 23)	–
γ, $\Delta\gamma$ (°)[a]	36, 46 (1, 9)	31, 60 (1, 1)	–
J, ΔJ (MHz)[a]	–	−3.2, 1.5 (0.6, 1.8)	–
η[b]	0.93	0.93	0.86

[a] The values in brackets determine the range in which 90% of the minimal RMSD is reached (Figure 5 and 8).
[b] This parameter was experimentally estimated to be roughly 0.9 by recording the frequency profile of the pump pulse (see Figure S10 in ref. [31]). During the fitting, η was varied in a small range around 0.9, [0.85, 0.95].

averaged over all $\Delta\nu$ (Figure 4(b)). Both distributions provide a similar value for the most probable distance; however, the distribution derived from the Tikhonov regularisation has a slightly larger width and contains artefacts at the distances of 1.5–1.6 nm due to the incomplete orientational averaging.

In order to investigate whether the optimised model reproduces the orientation selections for the different $\Delta\nu$, the $P(\theta) = \lambda(\theta)\sin\theta$ distribution was calculated for all frequency offsets and compared to the experimental dipolar spectra (Figure 4(c) and 4(d)). All calculated $P(\theta)$ distributions deviate from the $\sin\theta$ distribution indicating the presence of orientation selection. Moreover, all of them have a well-defined maximum, whose position depends on $\Delta\nu$. For the ZZ and YZ offsets, the $P(\theta)$ function has a highest probability at θ angles around 90°, whereas θ angles around 0° are deselected. The opposite case is found for the XX and YX offsets both having a highest probability at small θ angles. For the YY and XZ offsets, the intensities of both singularities are smaller than compared to the $\sin\theta$ function, whereas intermediate angles increase in probability. Comparing these distributions with the intensities of the parallel ($\nu_{\|}$, $\theta = 0°$) and perpendicular ($\nu_{\perp}$, $\theta = 90°$) frequency components, Equation (1) in the experimental Fourier transformed spectra reveals their consistency. Indeed, the dipolar spectra of the ZZ and YZ time traces contain only $\nu_{\perp}$, the dipolar spectra of the ZZ and YZ time traces is dominated by $\nu_{\|}$, and both the components determine the dipolar spectra of the YY and XZ time traces. Thus, one can conclude that the found geometric model is able to reproduce the orientation selections which are present in the PELDOR data.

Furthermore, the error estimate of each obtained geometric parameter is of interest. Since the optimisation procedure deals with a large number of fitting parameters (12 in this case) and they are interdependent, such error estimation is not trivial. Nevertheless, in order to provide some quantitative estimation of how defined the obtained parameters are, RMSD surfaces were recorded for the mean value and distribution width of each geometric parameter. Then, the parameter ranges in which 90% of the minimal RMSD is reached were determined. These ranges can be considered as a rough estimate of the confidence intervals for the corresponding parameters. While recording the RMSD surfaces for a certain pair of parameters, all other parameters were set to their optimised values. Note that it would be more robust to optimise all other parameters too, but this task is too time consuming. The RMSD surfaces recorded for **1** are shown in Figure 5, and the corresponding confidence intervals are given in Table 2 (in brackets). The RMSD surface recorded for μ and σ shows only one prominent minimum appearing at the distance, which was also found directly from the fitting of the PELDOR data. All other RMSD surfaces demonstrate that the angular parameters of the model are defined with a varying precision. This could be due to the different selectivity of the PELDOR measurements with respect to the different angles and to the overlapping of symmetry-related solutions.

Due to the inversion symmetry of the spin Hamiltonian, the obtained geometric parameters are not unique and several symmetry-related sets of parameters can provide a similar fit to the same PELDOR data-set [14]. Therefore, we examined whether the fits are sensitive to the 180° rotation of spin A and/or spin B around one of the g principal axes. This rotation influences only the ξ, φ, α, β, and γ angular parameters of the model, whereas the other parameters remain unchanged. In total, 16 symmetry-related sets of geometric parameters were found and then scored to provide the goodness of fit (RMSD). The results listed in Table 3 reveal a slight variation of the RMSD for the different solutions, which was also observed in Ref. [32]. Plotting the corresponding fits showed that only eight solutions with the RMSD ≤ 0.039 provide a good fit to the experimental time traces (Figure S1 in the Supporting Material). Moreover, since the A and B spins are indistinguishable for the nitroxide biradical **1**, all found solutions might represent a transformation from spin A to spin B or vice versa.

To summarise, a variety of the symmetry-related solutions introduces some uncertainty into the relative orientation of the spin centres, but it does not change the parameters of the distance distribution. Note that the current analysis aims to provide a model, which reproduces

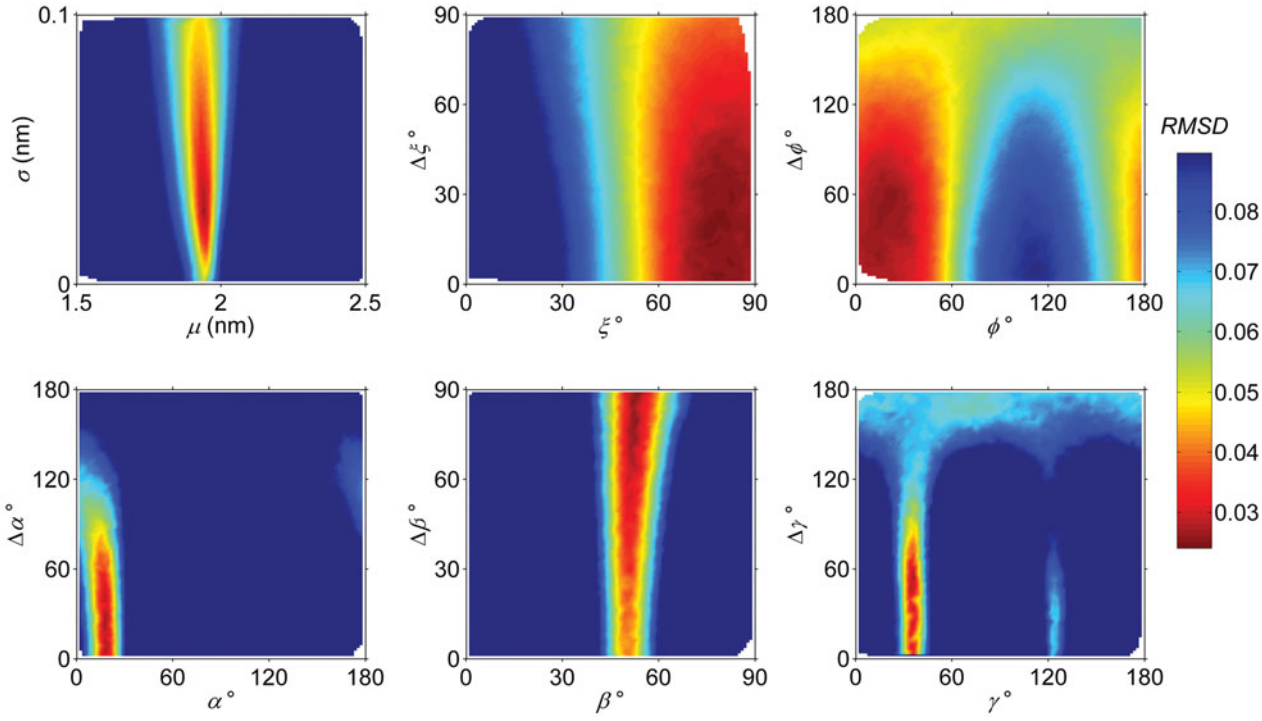


Figure 5. The RMSD surfaces for the geometric parameters, which were used to fit the PELDOR time traces of **1**. Each RMSD surface is recorded as a function of the mean value and distribution width of one of the geometric parameters. During the recording of each individual RMSD surface for a pair of the geometric parameters, all other parameters were set to their optimised values listed in Table 2.

orientation selections for the different $\Delta\nu$ and allows for an accurate estimation of a distance distribution rather than a sophisticated model of the whole system. If additional structural information about the spin system were available, this model could be extended to a more specific molecular model, e.g. as was done for biradical **1** in Ref. [30]. Nevertheless, the geometric model of **1** found here and the more elaborate model from the previous study should not exclude each other. In order to check this, the molecular model from Ref. [30] is reproduced in Figure 6(a): each nitroxide including the ester groups and the connecting bridge were represented by four serially connected vectors; the two nitroxide groups were allowed to rotate around the phenolic bond with the N–O bond (g_{xx} principle axis) tracing out

Table 3. Summary of the symmetry-related geometric models and their fitness to the PELDOR data of **1**.

Transformation[a]	ξ (°)	φ (°)	α (°)	β (°)	γ (°)	RMSD
Fitting result	76	25	19	52	36	0.025
Inversion of g_{xx}^{A}	104	335	19	52	36	0.059
Inversion of g_{yy}^{A}	104	155	19	52	36	0.054
Inversion of g_{zz}^{A}	76	205	19	52	36	0.039
Inversion of g_{xx}^{B}	76	25	199	128	144	0.025
Inversion of g_{yy}^{B}	76	25	199	128	324	0.025
Inversion of g_{zz}^{B}	76	25	19	52	216	0.024
Inversion of g_{xx}^{A} and g_{xx}^{B}	104	335	199	128	144	0.059
Inversion of g_{xx}^{A} and g_{yy}^{B}	104	335	199	128	324	0.059
Inversion of g_{xx}^{A} and g_{zz}^{B}	104	335	19	52	216	0.059
Inversion of g_{yy}^{A} and g_{xx}^{B}	104	155	199	128	144	0.053
Inversion of g_{yy}^{A} and g_{yy}^{B}	104	155	199	128	324	0.054
Inversion of g_{yy}^{A} and g_{zz}^{B}	104	155	19	52	216	0.054
Inversion of g_{zz}^{A} and g_{xx}^{B}	76	205	199	128	144	0.038
Inversion of g_{zz}^{A} and g_{yy}^{B}	76	205	199	128	324	0.038
Inversion of g_{zz}^{A} and g_{zz}^{B}	76	205	19	52	216	0.038

[a] g_{xx}^{A}, g_{yy}^{A}, and g_{zz}^{A} denote the principle components of the g-tensor of spin A; g_{xx}^{B}, g_{yy}^{B}, and g_{zz}^{B} denote the g-tensor for spin B.

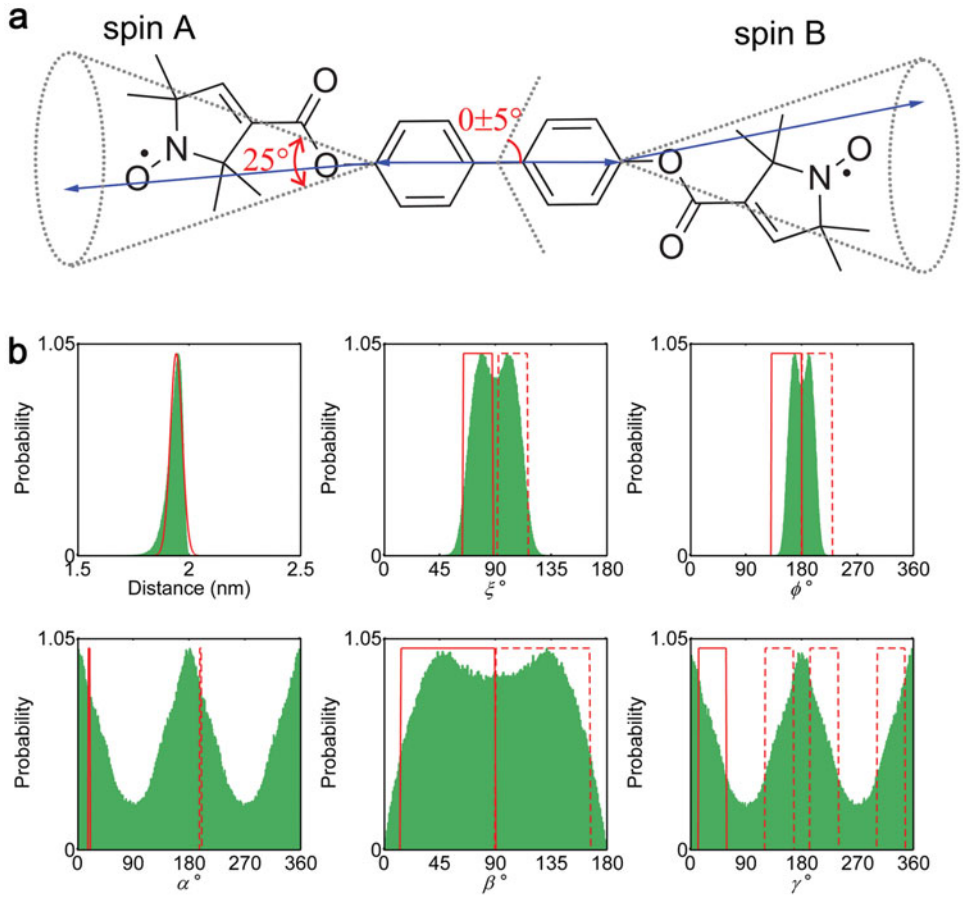


Figure 6. Comparison of the geometric models for biradical **1**. (a) The structure of **1** and its model as used in Ref. [30]. The model consists of four serially connected vectors (blue arrows): two vectors represent the nitroxide moieties including the ester groups, and the other two vectors represent the connecting bridge. The molecular flexibility is described by the backbone bending angle and the cone angle. (b) The distributions of the geometric parameters as determined in Figure 3 are calculated for the model described above (green). These distributions are overlaid with the corresponding distributions obtained from the fitting (red). The symmetry-related sets of the geometric parameters are shown by dashed lines.

a cone with a mean opening of 25°; and the flexibility of the connecting bridge was modelled by a bending motion of 5° about the biphenyl bond. Then, the distributions of the r, ξ, φ, α, β, and γ geometric parameters are calculated for this model and compared to the corresponding distributions obtained with the fitting algorithm used here (Figure 6(b)). In order to take into account the symmetry considerations for the geometric model, a superposition of the found symmetry-related solutions is included. As one can see, both the models predict the ξ angle to be distributed around 90° and the φ angle around 180°. Also, the widths of their distributions are very similar. The mean values of the Euler angles defining the relative orientations of two nitroxide g-tensors are close to each other too. Moreover, both the models predict a broad distribution of the β angle that agrees with the free rotation of nitroxide groups around the backbone of the molecule. At the same time, the shapes of α and γ distributions differ. This difference should not be over interpreted, since the RMSD surfaces reveal that the uncertainty of the obtained distribution widths is large, especially for α (Figure 5 and Table 2). This is the highly relevant, remembering that the previous model relied on simulations of the data with their goodness judged by eye. Nevertheless, the overall comparison of the two models shows that the simplified geometric model obtained from the fitting has essential similarities to the molecular model proposed in Ref. [30].

4.2. *Nitroxide biradical 2*

According to the previous studies [27,30], bisnitroxide **2** is more rigid than **1** and is characterised by an exchange coupling between the two nitroxide spins. The presence of exchange coupling is revealed by $\nu_{\parallel} \neq 2\nu_{\perp}$ (Figure 7(c)). Thus, reading out the experimental values of $\nu_{\parallel}$ and $\nu_{\perp}$ from the Fourier transformed spectra and using the equation $J = (\nu_{\parallel} + 2\nu_{\perp})/3$ [27], two possible estimates of the exchange coupling constant J were deduced, –3.3 and

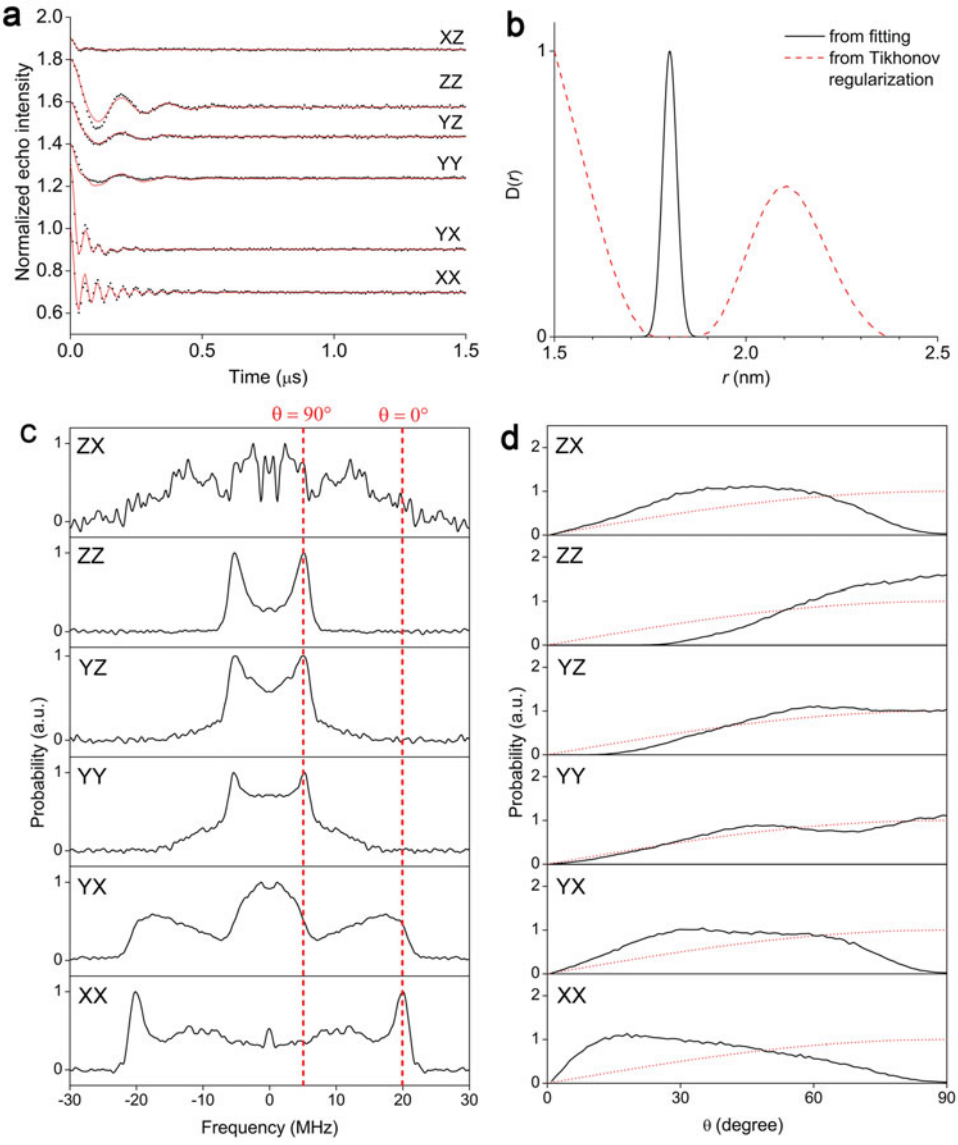


Figure 7. Fitting results for the PELDOR time traces of **2**. (a) The experimental time traces (dots) overlaid with their fits (lines). (b) The inter-spin distance distributions obtained from the fitting (solid line) and from the Tikhonov regularisation of the time trace averaged over all offsets (dash line). (c) The dipolar spectra obtained after Fourier transformation of each of the six experimental time traces. The positions of the parallel and perpendicular components are marked by red dashed lines. (d) The form factors $P(\theta)$ calculated for the optimised geometric model are shown for each time trace. A $\sin\theta$ distribution is depicted by a red dashed line and corresponds to the case without orientation selection.

–10.1 MHz. In order to obtain a more precise value of the exchange coupling constant, it was included in the fitting of PELDOR time traces in the form of two additional fitting parameters, the mean value J and its uniform distribution width ΔJ. The exchange coupling constant was taken into account in the current analysis by adding it to ω_{dd}, while the integral (4) was calculated.

The results of the fitting are presented in Figure 7. Figure 7(a) shows that again good fits to the experimental data were found for the parameters listed in Table 2. The obtained distance distribution with a mean value of 1.80 nm and a standard deviation of 0.02 nm is in good agreement with the previous data analysis [30]. The optimised value of exchange coupling constant J of –3.2 MHz and the distribution width ΔJ of ±0.75 MHz are also very close to the one found in the previous work [30]. At the same time, summing up all PELDOR time traces of **2** and analysing the resulting time trace with the DeerAnalysis program yields an incorrect distance distribution, because this program neglects the exchange coupling between spins (Figure 7(b)).

Figure 7(c) and 7(d) demonstrates that the orientation selections are well reproduced in the corresponding $P(\theta)$ dependences calculated for the optimised geometric model.

One can see from both the figures that the θ angles around 0° are dominant for $\Delta\nu$ involving the X component of the nitroxide *g*-tensor. The opposite tendency is observed for the ZZ offset whose dipolar spectrum is dominated by the dipolar frequencies corresponding to the θ angles around 90°. The YY and YZ offsets do not possess a significant orientation selection. Thus, these results show that, as in the case of biradical **1**, the optimised model of biradical **2** is able to provide orientation selections, which are consistent with the experimental ones.

In the second step, the RMSD surfaces for the geometric parameters were recorded and the error of each parameter was estimated in the same way as above (Figure 8(a)). The RMSD surface of the distance parameters μ and σ has a single minimum whose coordinates are in agreement with the distances obtained after the fitting. As in the case of biradical **1**, all angular parameters are determined with a different precision (Table 2). Interestingly, the confidence intervals of most of the angular parameters are smaller compared to the same values for molecule **1**. This can be attributed to the higher rigidity of the biradical **2** in comparison to **1**. In addition, the RMSD surface was recorded for the exchange coupling constant (Figure 8(b)). The minimum of this surface appears at $J = -3.2 \pm 0.3$ MHz and $\Delta J = 1.5 \pm 0.9$ MHz, in accordance with the fitting result. The second possible solution is clearly excluded.

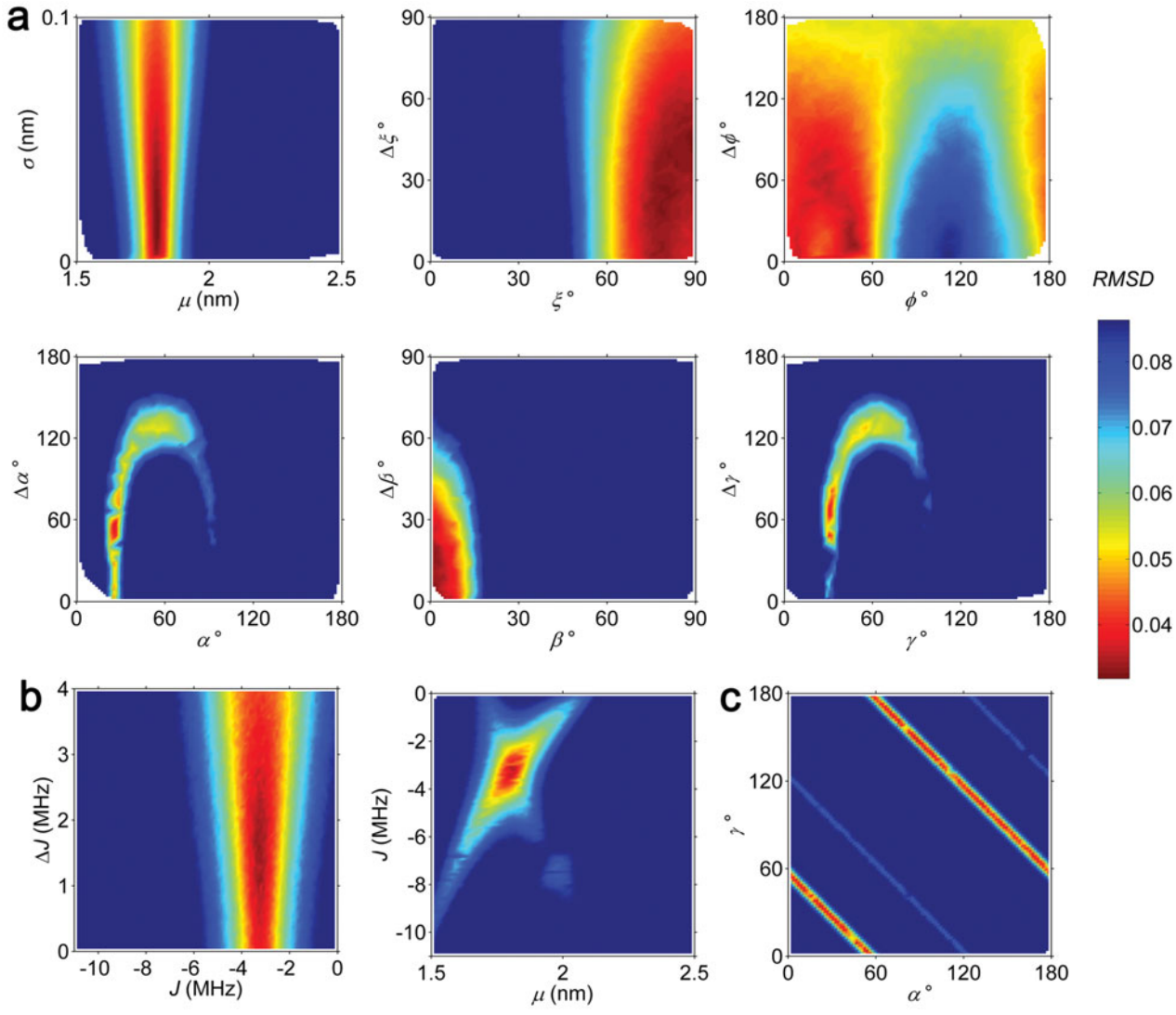


Figure 8. The RMSD surfaces for the parameters, which were used to fit the PELDOR time traces of **2**. (a) Each RMSD surface is recorded as a function of the mean value and distribution width of one of the geometric parameters. (b) The RMSD surfaces for the exchange coupling constant parameters. (c) The RMSD surface for the α and γ parameters. The global minimum appears at $(\alpha + \gamma) = 240°$. The symmetry-related solution corresponds to $(\alpha + \gamma) = 60°$. During the recording of each individual RMSD surface for a pair of fitting parameters, all other parameters were set to their optimised values listed in Table 2.

Table 4. Summary of the symmetry-related geometric models and their fitness to the PELDOR data of **2**.

Transformation[a]	ξ (°)	φ (°)	α (°)	β (°)	γ (°)	RMSD
Fitting result	79	46	26	3	31	0.032
Inversion of g_{xx}^{A}	101	314	26	3	31	0.076
Inversion of g_{yy}^{A}	101	134	26	3	31	0.076
Inversion of g_{zz}^{A}	79	226	26	3	31	0.033
Inversion of g_{xx}^{B}	79	46	206	177	149	0.032
Inversion of g_{yy}^{B}	79	46	206	177	329	0.032
Inversion of g_{zz}^{B}	79	46	26	3	211	0.032
Inversion of g_{xx}^{A} and g_{xx}^{B}	101	314	206	177	149	0.076
Inversion of g_{xx}^{A} and g_{yy}^{B}	101	314	206	177	329	0.076
Inversion of g_{xx}^{A} and g_{zz}^{B}	101	314	26	3	211	0.076
Inversion of g_{yy}^{A} and g_{xx}^{B}	101	134	206	177	149	0.076
Inversion of g_{yy}^{A} and g_{yy}^{B}	101	134	206	177	329	0.076
Inversion of g_{yy}^{A} and g_{zz}^{B}	101	134	26	3	211	0.076
Inversion of g_{zz}^{A} and g_{xx}^{B}	79	226	206	177	149	0.033
Inversion of g_{zz}^{A} and g_{yy}^{B}	79	226	206	177	329	0.033
Inversion of g_{zz}^{A} and g_{zz}^{B}	79	226	26	3	211	0.033

[a] g_{xx}^{A}, g_{yy}^{A}, and g_{zz}^{A} denote the principle components of the *g*-tensor of spin A; g_{xx}^{B}, g_{yy}^{B}, and g_{zz}^{B} denote the *g*-tensor for spin B.

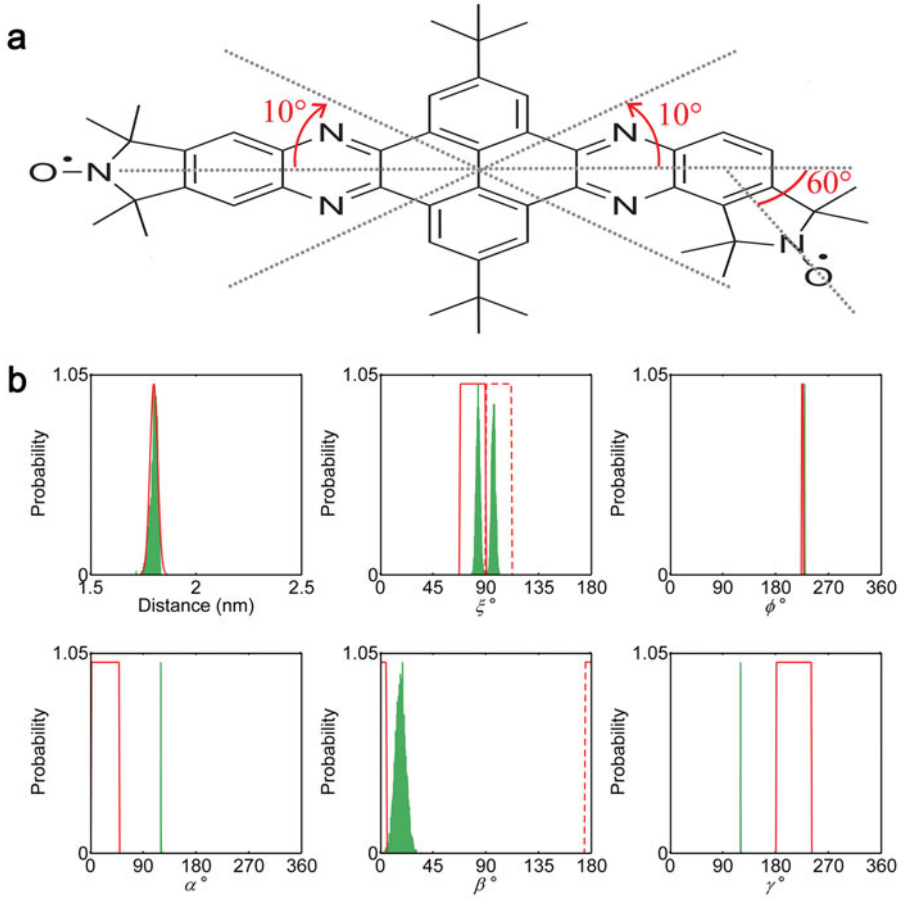


Figure 9. Comparison of the geometric models for biradical **2**. (a) The structure of **2** and its model as used in Ref. [30]. One N–O bond is collinear to the molecular backbone, and another N–O bond is turned from it by 60°. The molecular backbone is bended about the centre of the molecule by ±10° with a standard deviation of 2.5°. (b) The distributions of the geometric parameters as determined in Figure 3 are calculated for the model described above (green). These distributions are overlaid with the corresponding distributions obtained from the fitting (red). The relevant symmetry-related sets of the geometric parameters are shown by dashed lines.

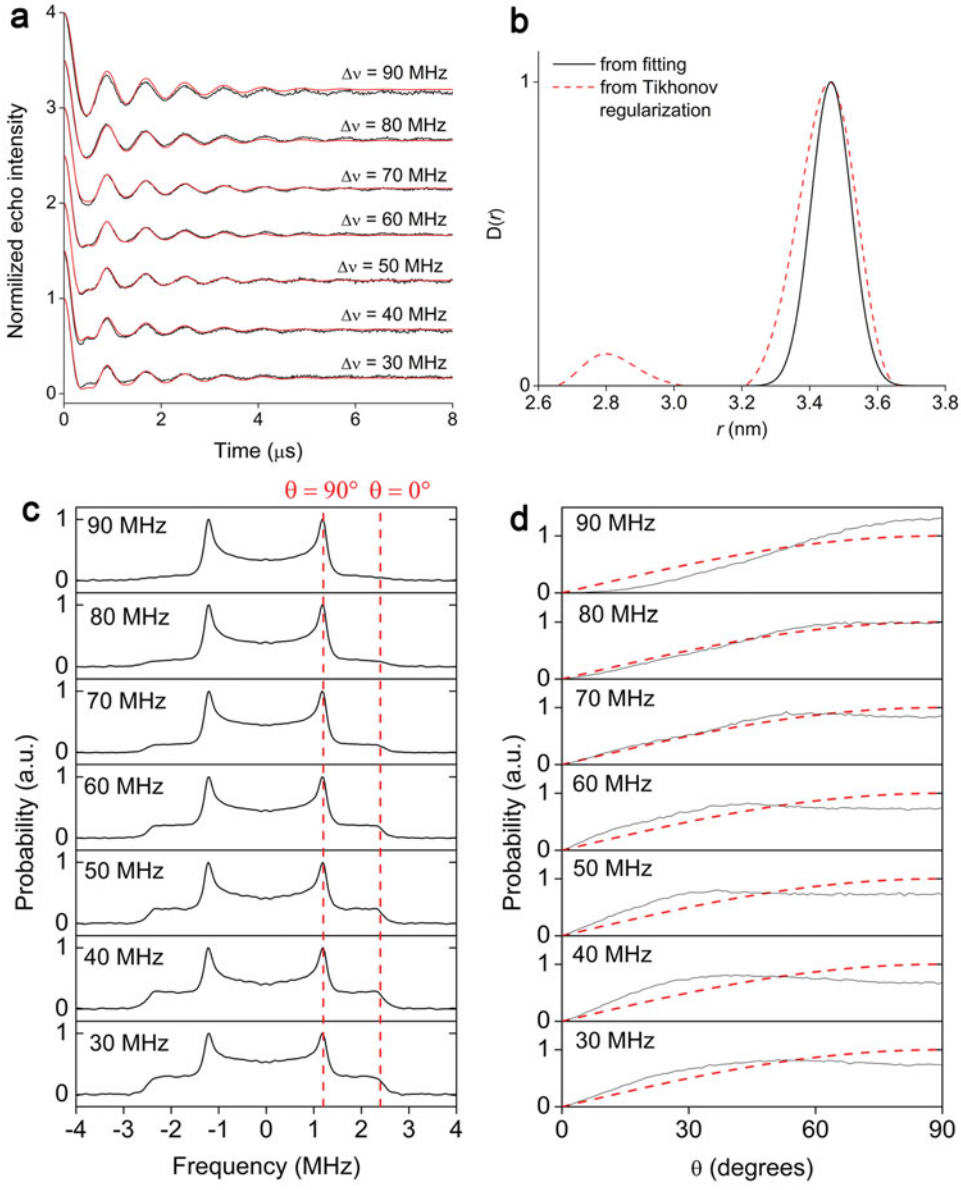


Figure 10. Fitting results for the PELDOR time traces of **3**. (a) The experimental time traces (dots) overlaid with their fits (lines). (b) The inter-spin distance distributions obtained from the fitting (solid line) and from the Tikhonov regularisation of the time trace averaged over all offsets (dashed line). (c) The dipolar spectra obtained after Fourier transformation of each of the six experimental time traces. The positions of the parallel and perpendicular components are marked by red dashed lines. (d) The form factors $P(\theta)$ calculated for the optimised geometric model are shown for each time trace. A $\sin\theta$ distribution is depicted by a red dashed line and corresponds to the case without orientation selection.

The same symmetry considerations as mentioned for **1** are also valid for the geometric model of **2**. The 180° rotation of spin A and/or spin B around one of the *g* principal axes provides 16 symmetry-related sets of angular parameters; however, only eight of them provide a good fit to the experimental time traces (Table 4, Figure S2 in the Supporting Material). Again, there are two possibilities of assigning two nitroxide spins to the A and B spins of the model used. This doubles the number of possible symmetry-related solutions. Note that the solution found from the fitting corresponds to the case when the spin of the right nitroxide group of **2** in Figure 1 is considered as spin A.

Finally, the geometric model of **2** obtained here is compared with the model proposed in Ref. [30]. The latter model is shown in Figure 9(a). In this model, one nitroxide is aligned such that the g_{xx} principle axis is parallel to molecular backbone, whereas the g_{xx} principle axis of the other nitroxide is rotated by 60° about its g_{zz} principle axis. The molecular backbone is bent about the centre of the molecule by ±10° with a standard deviation of 2.5°. In

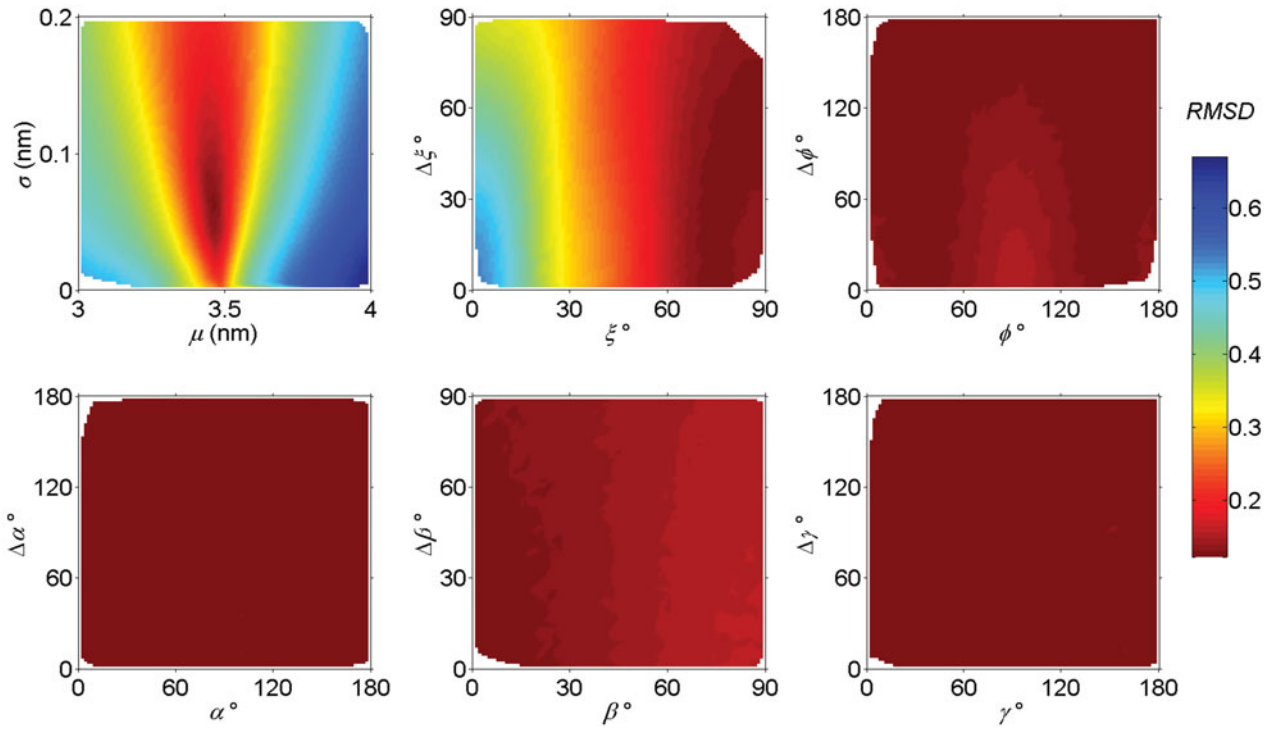


Figure 11. The RMSD surfaces for the geometric parameters, which were used to fit the PELDOR time traces of **3**. Each RMSD surface is recorded as a function of the mean value and distribution width of one of the geometric parameters. During the recording of each individual RMSD surface for a pair of the geometric parameters, other parameters were set to their optimised values listed in Table 2.

analogy to the previous paragraph, the distributions of the r, ξ, φ, α, β, and γ parameters were calculated and compared for both the models (Figure 9(b)). This comparison reveals that the value of the φ angle is identical for both the models. Moreover, both the models predict a slight bending of the molecular backbone, which results in distributions of the ξ and β angles. The values for the bending angle slightly differ. In the previous model, the mean bending angle was centred at $\pm 10^\circ$ shifting the mean values of ξ and β from 90° and 0°, respectively. The model used here yields the ξ and β distributions centred around 90° and 0°, respectively. This corresponds to a mean bending angle of 0°. A mean bending angle of 0° was also used in Ref. [27] in conjunction with a distribution of $\pm 5^\circ$. Note that this discrepancy in the bending angle could originate from the high uncertainty of the ξ and β parameters revealed by the corresponding RMSD surfaces (Figure 8(a)). The distributions of α and γ Euler angles differ for the two models. Nevertheless, the difference in the mean α and γ angles is related only to the features of the Euler angles within the chosen z–x'–z'' convention: when the β angle is close to 0°, the α and γ angles denote a rotation about the same z-axis and, therefore, all values of α and γ which provide the same sum $(\alpha + \gamma)$ are identical. Thus, only the sum $(\alpha + \gamma)$ should be compared for the two models. Indeed, for both the models this sum is equal to 240° (Figure 8(c)).

To conclude, the comparison of the simplified geometric model with the more elaborate molecular model reveals their similarities and proves that the ability of the geometric model to correctly predict the orientation selections of the PELDOR data is not a coincidence.

4.3. *Trityl-nitroxide biradical 3*

In the following, the X-band PELDOR data-set of the trityl-nitroxide biradical **3** was used as a test for a case when only a sub-set of the geometric parameters are used due to the spectroscopic properties of the spin centres. Due to the low spectral resolution of the X-band experiments and low anisotropy of the trityl g-tensor, this spin centre can be considered as isotropic. For the same reasons, the g- and A-tensors of the nitroxide spin ($g_{xx} = 2.0093$, $g_{yy} = 2.0059$, $g_{zz} = 2.0018$, $A_{xx} = A_{yy} = 18$ MHz, $A_{zz} = 93$ MHz [31]) have a symmetry, which is close to axial. As a consequence, the number of the geometric parameters required to analyse the PELDOR time traces of **3** can be reduced from 12 down to 4 (Table 1). Thus, the set of the fitting parameters consists of the mean distance μ, its standard deviation σ, and two parameters of the ξ distribution. Note that the spin density distribution of the trityl spin is neglected in the present analysis. This was shown to be a valid approximation in the previous work as well, because the scale of the spin

density distribution is considerably smaller than the spin–spin distance [31].

The fitting yielded good fits to all experimental time traces (Figure 10(a)). The optimised trityl-nitroxide distance distribution with $\mu = 3.46$ nm and $\sigma = 0.6$ nm is in agreement with the previous studies [31]. It is also similar to the distribution derived after the Tikhonov regularisation of the averaged time trace (Figure 10(b)), but the fitting approach provides a slightly different width of the distance distribution and also avoids the artefacts appearing around 2.8 nm. These artefacts may originate from incomplete orientation averaging over the nitroxide spectrum or from the incomplete suppression of the deuterium electron spin echo envelope modulation (ESEEM) in the PELDOR experiments [31].

The $P(\theta)$ dependences calculated for the optimised distributions of r and ξ are consistent with the corresponding dipolar spectra (Figure 10(c) and 10(d)). The maximum of the $P(\theta)$ dependence shifts from the small θ angles towards 90° with increasing $\Delta\nu$. According to that, the parallel component of the dipolar spectra decreases in intensity and the perpendicular component raises in intensity with the increasing $\Delta\nu$.

To check whether the reduced set of fitting parameters is sufficient for the present analysis, the fitting of the PELDOR time traces was repeated with the whole set of 12 geometric parameters. The fits with the same RMSD and geometric parameters as before were obtained (Figure S3 in the Supporting Material). In order to demonstrate the sensitivity of the fitting procedure with respect to each individual geometric parameter, the RMSD surfaces were recorded as described above (Figure 11). The RMSD surfaces clearly show that only one angular parameter ξ and its distribution $\Delta\xi$ are important for describing the PELDOR time traces of **3**.

The high uncertainty of the geometric parameters does not allow the construction of a geometric model of the spin pair here. Moreover, our fitting results reveal that the structural model used in the previous work to simulate the PELDOR time traces of **3** [31] can be only considered as one of several possible solutions.

5. Conclusion

A fitting algorithm of orientation-selective PELDOR data analysis was introduced and successfully tested. According to our knowledge, the genetic algorithm is applied for the first time to this problem. Analysis of the fitting results allows us to presume that the genetic algorithm found the global minimum of the fitting problems considered here. The algorithm finds a solution within hours. The global fitting of six time traces of **1** with 500 optimisation cycles took 4.84 hours, the same procedure for **2** lasted 3.94 hours. 10.08 hours were required to obtain the fits of the seven time traces of **3**.

For all examined PELDOR data-sets, reasonable fits of the experimental time traces were obtained. The optimised distance distributions are in agreement with the predictions made in the previous studies. This proves that the present algorithm is able to provide an accurate estimate of the mean value and width of the inter-spin distance distribution. For biradicals **1** and **3**, the algorithm requires a smaller number of time traces for recovering the proper distance distribution than the procedure, which includes averaging the time traces over different frequency offsets and analysing it, using the DeerAnalysis program. This feature could be important when not all spectral components can be probed due to the duration of the experiment and/or restrictions imposed by the spectrometer. In addition, the latter procedure loses the orientation information.

Within the point dipole approximation, this approach allows for the analysis of PELDOR data without a priori knowledge about the structure and dynamics of the molecular system. The simplified geometric model of the spin pair is described by 12 parameters plus 2 parameters for the exchange coupling if present, and all are optimised during the fitting. Despite all simplifications made for the model, our tests show its ability to reproduce the orientation selections present in the experimental PELDOR data. Moreover, the obtained models are comparable to the previous models with more elaborate modes of motions. This shows that neglecting the correlation between individual geometric parameters is a viable approximation for the molecules studied here. However, this does not mean this assumption is suitable for all possible molecular systems, especially for those, where such correlations are strong.

Acknowledgements

Olav Schiemann gratefully acknowledges financial support of the Deutsche Forschungsgemeinschaft via the projects A6 and C8 within the Sonderforschungsbereich SFB813 ‘Chemistry at Spin Centers’. Graham M. Smith would like to acknowledge funding from EPSRC [grant number F004583/1] and the UK Research Councils under the Basic Technology Program.

Supplemental data

Supplemental data for this article can be accessed here.

References

[1] A.D. Milov, K.M. Salikhov, and M.D. Shirov, Sov. Phys. Solid State **23**, 565 (1981).
[2] R.E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, and H.W. Spiess, Angew. Chem. Int. Ed. **37**, 2833 (1998).
[3] P.P. Borbat and J.H. Freed, in *Structure and Bonding*, edited by C.R. Timmel and J.R. Harmer (Springer-Verlag, Berlin Heidelberg, 2014), vol. 152, pp. 1–82.
[4] G. Jeschke, Annu. Rev. Phys. Chem. **63**, 419 (2012).
[5] R. Ward and O. Schiemann, in *Structure and Bonding*, edited by C.R. Timmel and J.R. Harmer (Springer-Verlag, Berlin Heidelberg, 2014), vol. 152, pp. 249–282.

[6] J. Klare and H.-J. Steinhoff, in *Encyclopedia of Analytical Chemistry*, edited by R.A. Meyers (John Wiley, Chichester, 2010).
[7] I. Krstic, A. Marko, C.M. Grytz, B. Endeward, and T. F. Prisner, in *RNA Structure and Folding*, edited by D. Klostermeier and Ch. Hammann (DE GRUYTER, Berlin, 2013).
[8] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C.R. Timmel, D. Hilger, and H. Jung, Appl. Magn. Reson. **30**, 473 (2006).
[9] V.P. Denysenkov, T.F. Prisner, J. Stubbe, and M. Bennati, Proc. Natl. Acad. Sci. U.S.A. **103**, 13386 (2006).
[10] V. Denysenkov, D. Biglino, W. Lubitz, T. Prisner, and M. Bennati, Angew. Chem. Int. Ed. **47**, 1224 (2008).
[11] O. Schiemann, P. Cekan, D. Margraf, T.F. Prisner, and S.T. Sigurdsson, Angew. Chem. Int. Ed. **48**, 3292 (2009).
[12] A. Marko, V.P. Denysenkov, D. Margraf, P. Cekan, O. Schiemann, S.T. Sigurdsson, and T.F. Prisner, J. Am. Chem. Soc. **133**, 13375 (2011).
[13] B. Endeward, J.A. Butterwick, R. MacKinnon, and T.F. Prisner, J. Am. Chem. Soc. **131**, 15246 (2009).
[14] C. Abé, D. Klose, F. Dietrich, W.H. Ziegler, Y. Polyhach, G. Jeschke, and H.-J. Steinhoff, J. Magn. Reson. **216**, 53 (2012).
[15] A.M. Bowen, C.E. Tait, C.R. Timmel, and J.R. Harmer, in *Structure and Bonding*, edited by C.R. Timmel and J.R. Harmer (Springer-Verlag, Berlin Heidelberg, 2014), vol. 152. pp. 283–328.
[16] I.M.C. van Amsterdam, M. Ubbink, G.W. Canters, and M. Huber, Angew. Chem. Int. Ed. **42**, 62 (2003).
[17] B.E. Bode, J. Plackmeyer, T.F. Prisner, and O. Schiemann, J. Phys. Chem. A **112**, 5064 (2008).
[18] Z. Yang, D. Kise, and S. Saxena, J. Phys. Chem. B **114**, 6165 (2010).
[19] Z. Yang, M.R. Kurpiewski, M. Ji, J.E. Townsend, P. Mehta, L. Jen-Jacobson, and S. Saxena, Proc. Natl. Acad. Sci. U.S.A. **109**, E993 (2012).
[20] C. Elsässer , M. Brecht, and R. Bittl, J. Am. Chem. Soc. **124**, 12606 (2002).
[21] J.E. Lovett, A.M. Bowen, C.R. Timmel, M.W. Jones, J.R. Dilworth, D. Caprotti, S.G. Bell, L.L. Wong, and J. Harmer, Phys. Chem. Chem. Phys. **11**, 6840 (2009).
[22] M.M. Roessler, M.S. King, A.J. Robinson, F.A. Armstrong, J. Harmer, and J. Hirst, Proc. Natl. Acad. Sci. U.S.A. **107**, 1930 (2010).
[23] S. Nakazawa, S. Nishida, T. Ise, T. Yoshino, N. Mori, R.D. Rahimi, K. Sato, Y. Morita, K. Toyota, D. Shiomi, M. Kitagawa, H. Hara, P. Carl, P. Höfer, and T. Takui, Angew. Chem. Int. Ed. **51**, 9860 (2012).
[24] A. Godt, M. Schulte, H. Zimmermann, and G. Jeschke, Angew. Chem. Int. Ed. **45**, 7560 (2006).
[25] A. Godt, M. Schulte, H. Zimmermann, and G. Jeschke, Angew. Chem. Int. Ed. **118**, 7722 (2006).
[26] D. Margraf, B.E. Bode, A. Marko, O. Schiemann, and T.F. Prisner, Mol. Phys. **105**, 2153 (2007).
[27] D. Margraf, P. Cekan, T.F. Prisner, S.Th. Sigurdsson, and O. Schiemann, Phys. Chem. Chem. Phys. **11**, 6708 (2009).
[28] A. Savitsky, A. Dubinskii, H. Zimmermann, W. Lubitz, and K. Moebius, J. Phys. Chem. B **115**, 11950 (2011).
[29] G.W. Reginsson and O. Schiemann, Biochem. J. **434**, 353 (2011).
[30] G.W. Reginsson, R.I. Hunter, P.A.S. Cruickshank, D.R. Bolton, S.T. Sigurdsson, G.M. Smith, and O. Schiemann, J. Magn. Reson. **216**, 175 (2012).
[31] G.W. Reginsson, N.C. Kunjir, S.Th. Sigurdsson, and O. Schiemann, Chem. Eur. J. **18**, 13580 (2012).
[32] I. Tkach, S. Pornsuwan, C. Hoebartner, F. Wachowius, S.Th. Sigurdsson, T.Y. Baranova, U. Diederichsen, G. Sicolia, and M. Bennati, Phys. Chem. Chem. Phys. **15**, 3433 (2013).
[33] A. Marko and T.F. Prisner, Phys. Chem. Chem. Phys. **15**, 619 (2013).
[34] B. Filipic and J. Strancar, Appl. Soft Comput. **1**, 83 (2001).
[35] T. Spalek, P. Pietrzyk, and Z. Sojka, J. Chem. Inf. Model. **45**, 18 (2005).
[36] S. Stoll and A. Schweiger, J. Magn. Reson. **178**, 42 (2006).
[37] A. Weber, O. Schiemann, B. Bode, and T.F. Prisner, J. Magn. Reson. **157**, 277 (2002).
[38] P.A. Cruickshank, D.R. Bolton, D.A. Robertson, R.I. Hunter, R.J. Wylde, and G.M. Smith, Rev. Sci. Instrum. **80**, 103102 (2009).
[39] D. Goldfarb, in *Structure and Bonding*, edited by C.R. Timmel and J.R. Harmer (Springer-Verlag, Berlin Heidelberg, 2014), vol. 152, pp. 163–204.
[40] A. Marko, D. Margraf, H. Yu, Y. Mu, G. Stock, and T.F. Prisner, J. Chem. Phys. **130**, 064102 (2009).
[41] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley Longman Publishing Co., Boston, 1989).

Geometric model based fitting algorithm of orientation selective PELDOR data

Dinar Abdullin[a], Gregor Hagelueken[a], Robert I. Hunter[b], Graham M. Smith[b], Olav Schiemann[a,*]

[a]*Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany.*
[b]*School of Physics, University of St Andrews, North Haugh, KY16 9ST St Andrews, UK.*
**Author to whom correspondence should be addressed: schiemann@pc.uni-bonn.de, phone: +49(0)228-732989.*

Supporting Material

Content

1. Input data for the fitting of PELDOR data of 1, 2, and 3

As it was mentioned in the main text the fitting algorithm uses the following input data: 1) spectroscopic parameters of the spin centers, 2) experimental settings of the PELDOR measurements, and 3) background corrected PELDOR time traces.

Spectroscopic parameters of the spin centers of **1**-**3** were taken from the literature [30, 31]. The parameters which were used in the fitting of the PELDOR time traces of **1**-**3** are listed in the Table S1. Note that the hyperfine couplings of the trityl spin to the carbon nuclei A(^{13}C) were neglected due to the low abundance of the ^{13}C.

The detailed explanation of PELDOR measurements on **1**-**3** can be found in ref. [30, 31]. The experimental settings which were used in the fitting are listed in the Table S2.

All PELDOR time traces of **1**-**3** are shown in the main text. The background correction of the experimental time traces was done by means of the DeerAnalysis program.

Table S1. Spectroscopic parameters used for the fitting of the PELDOR time traces of biradicals **1**, **2**, and **3**.

Spectroscopic	**1**	**2**	**3**	
parameters	Nitroxide	Nitroxide	Nitroxide	Trityl
g_{xx}, g_{yy}, g_{zz}	2.0104, 2.0073, 2.0033	2.0100, 2.0072, 2.0033	2.0093, 2.0059, 2.0018	2.0030, 2.0027, 2.0021
g-strain	0.0004, 0.0003, 0.0001	0.005, 0.0002, 0.0003	-	-
A_{xx}, A_{yy}, A_{zz} (MHz)				
A(^{14}N)	8, 6, 96	10, 12, 90	18, 18, 93	-
A-strain (MHz)	0, 0, 12	5, 0, 5	-	-
Linewidth (MHz)[c]	22.4	11.0	14.0	3.64

[a] The spectroscopic parameters are adopted from ref. [29]. [b] The spectroscopic parameters are adopted from ref. [31]. [c] The linewidth is given as a peak-to-peak linewidth of the Gaussian shaped line.

Table S2. The experimental PELDOR parameters of **1**, **2**, and **3**.

Molecule	$\Delta\nu$	ν_{det} (GHz)	ν_{pump} (GHz)	B_0 (T)	$t_{\pi/2}$ (ns)	t_{π} (ns)	t_{pump} (ns)
1	XX	93.9996	93.9300	3.3415	7	14	14
	YY	93.9996	93.9300	3.3443	7	14	14
	ZZ	93.9996	94.1004	3.3515	7	14	16
	YX	93.9996	93.8880	3.3415	7	14	18
	YZ	93.9996	94.1004	3.3490	7	14	14
	ZX	93.9996	93.7896	3.3415	7	14	23
2	XX	93.9996	93.9300	3.3415	8	16	16
	YY	93.9996	93.9300	3.3443	8	16	24
	ZZ	93.9996	94.1004	3.3515	8	16	17
	YX	93.9996	93.8880	3.3415	8	16	16
	YZ	93.9996	94.1004	3.3490	8	16	16
	ZX	93.9996	93.7896	3.3415	8	16	15
3	30 MHz	9.751631	9.721578	0.347010	16	32	16
	40 MHz	9.761410	9.721578	0.347010	16	32	16
	50 MHz	9.771577	9.721578	0.347010	16	32	16
	60 MHz	9.786335	9.726154	0.347170	16	32	16
	70 MHz	9.795081	9.721931	0.347000	16	32	16
	80 MHz	9.805863	9.726154	0.347170	16	32	16
	90 MHz	9.813152	9.723332	0.347070	16	32	32

2. PELDOR fits for biradical 1

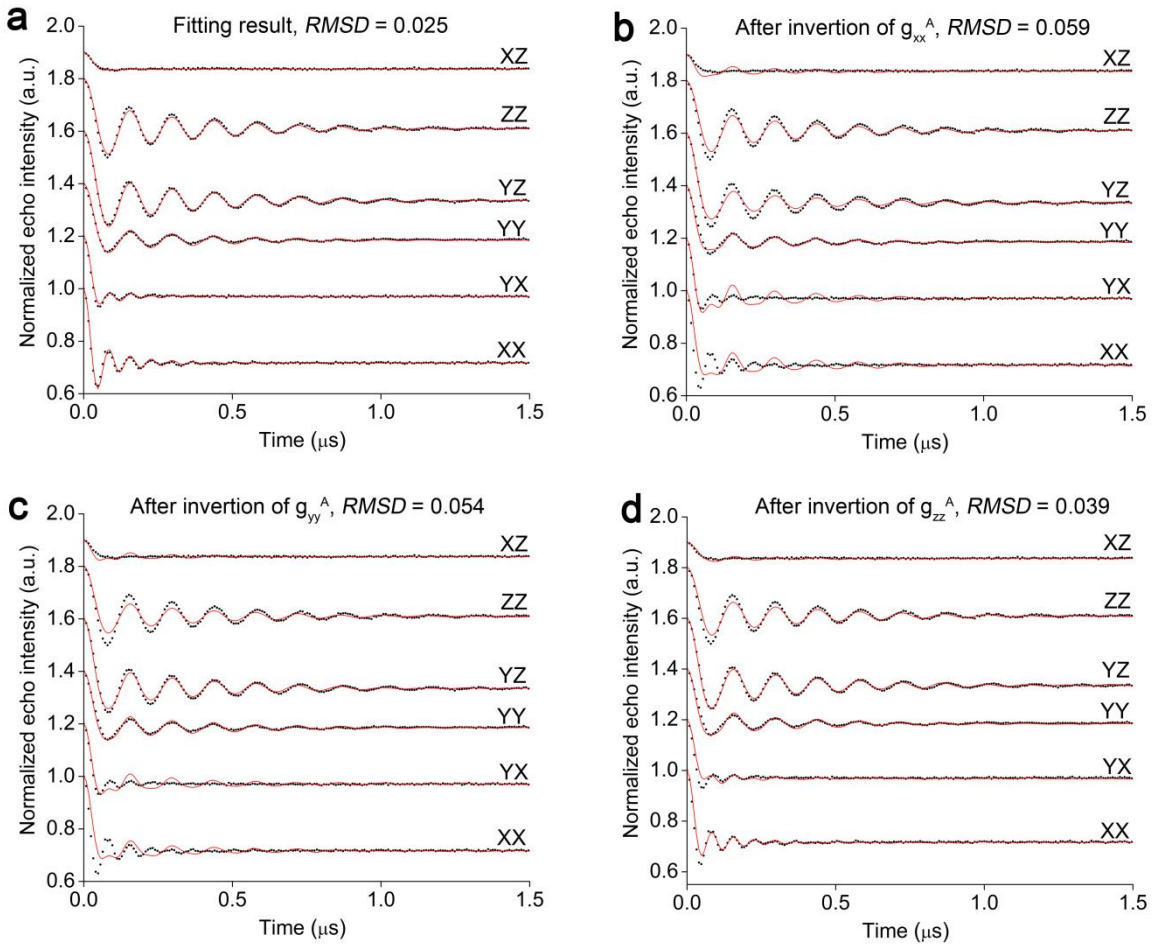


Figure S1. The fits of the PELDOR time traces of **1** corresponding to (a) the optimized set of geometric parameters (Table 2) and the symmetry-related sets of parameters (Table 3) obtained after the inversion of (b) the g_{xx} principle axis of spin A, (c) g_{yy} principle axis of spin A, and (d) g_{zz} principle axis of the spin A.

3. PELDOR fits for biradical 2

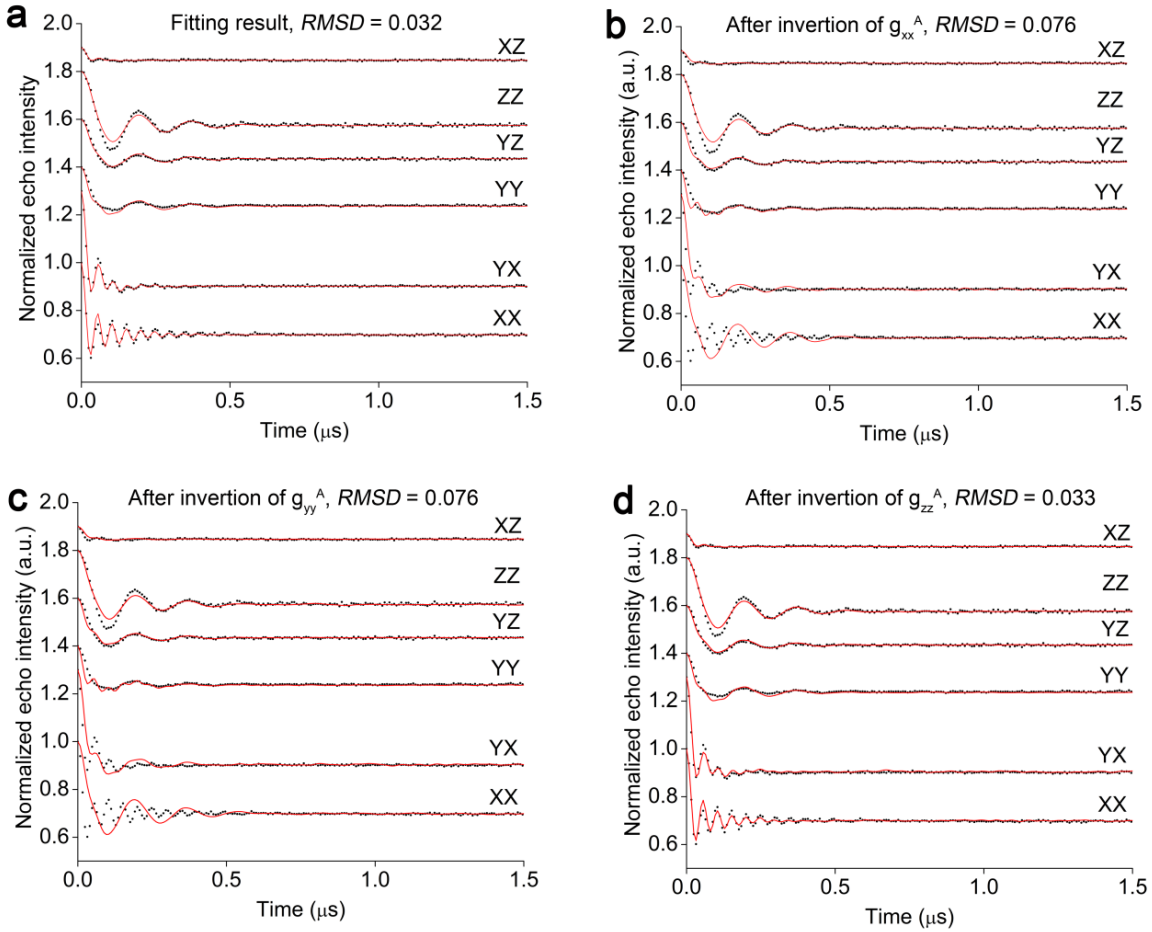


Figure S2. The fits of the PELDOR time traces of **2** corresponding to (a) the optimized set of geometric parameters (Table 2) and the symmetry-related sets of parameters (Table 4) obtained after the inversion of (b) the g_{xx} principle axis of spin A, (c) g_{yy} principle axis of spin A, and (d) g_{zz} principle axis of the spin A.

4. PELDOR fits for biradical 3

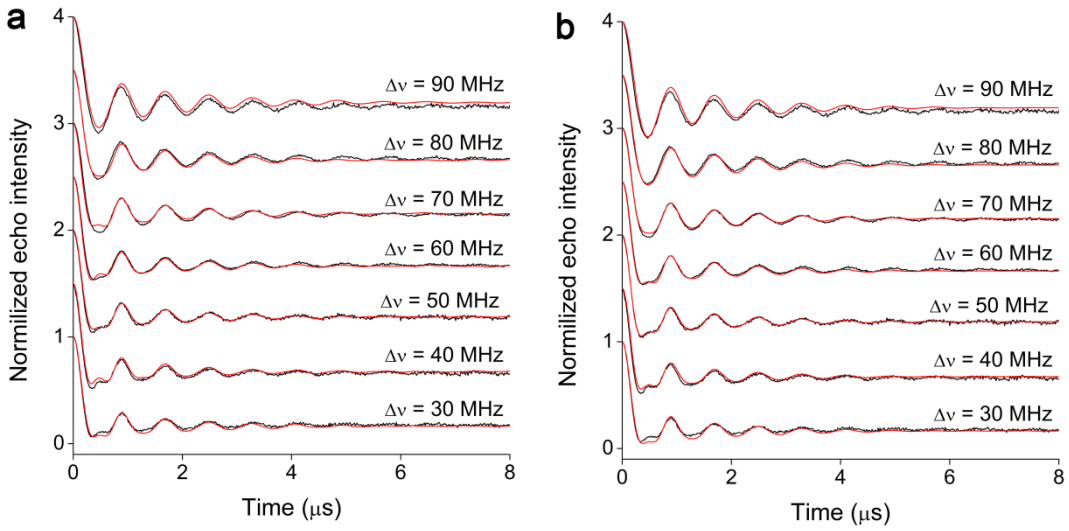


Figure S3. The fits of the PELDOR time traces of **3** obtained after optimization of (a) 4 and (b) 12 geometric parameters.

[P2] mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL

Reprinted with permission from

Gregor Hagelueken, Dinar Abdullin, Richard Ward, Olav Schiemann, "mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL", *Mol. Phys.* **2013**, *111*, 2757–2766. DOI: 10.1080/00268976.2013.809804

Own contribution to the manuscript:

- Development and testing of the program mtsslTrilaterate.

Molecular Physics, 2013
Vol. 111, Nos. 18–19, 2757–2766, http://dx.doi.org/10.1080/00268976.2013.809804

Taylor & Francis
Taylor & Francis Group

INVITED ARTICLE

mtsslSuite: *In silico* spin labelling, trilateration and distance-constrained rigid body docking in PyMOL

Gregor Hagelueken[a,*], Dinar Abdullin[a], Richard Ward[b] and Olav Schiemann[a,b]

[a]*Institute of Physical and Theoretical Chemistry, University of Bonn, Bonn, Germany;* [b]*Biomedical Sciences Research Complex, The University of St. Andrews, Fife, UK*

(*Received 8 March 2013; final version received 18 May 2013*)

Nanometer distance measurements based on electron paramagnetic resonance methods in combination with site-directed spin labelling are powerful tools for the structural analysis of macromolecules. The software package mtsslSuite provides scientists with a set of tools for the translation of experimental distance distributions into structural information. The package is based on the previously published mtsslWizard software for *in silico* spin labelling. The mtsslSuite includes a new version of MtsslWizard that has improved performance and now includes additional types of spin labels. Moreover, it contains applications for the trilateration of paramagnetic centres in biomolecules and for rigid-body docking of subdomains of macromolecular complexes. The mtsslSuite is tested on a number of challenging test cases and its strengths and weaknesses are evaluated.

Keywords: DEER; PELDOR; triangulation; EPR; FRET; docking

Introduction

Electron paramagnetic resonance (EPR) based distance measurements in biological macromolecules are a valuable complement to crystallography or NMR studies [1–3]. The most commonly used EPR technique for distance measurements is pulsed electron–electron double resonance (PELDOR, also known as DEER) [4]. It can be used to accurately determine interspin distances in the range within 15–80 Å [5]. Since most biological macromolecules are not paramagnetic, site-directed spin labelling (SDSL) with nitroxides is often used to incorporate spins into proteins [6,7] or nucleic acids [8,9]. However, the translation of PELDOR-derived distances into structures is often affected by the flexibility and dimension of the spin labels used. Thus, an accurate model of the conformational distribution of the spin labels attached to the biomolecule is required to draw structural conclusions from the distance data. Since crystallographic studies of multiple spin-labelled mutants of a biomolecule are in general not feasible; this model will in most cases be based on computer simulations. The currently available simulation programs MMM, mtsslWizard and PRONOX that generate such label distributions predict experimental PELDOR distances between two nitroxide spin labels with errors in the range of ∼3 Å [10–12].

Once, a model of the macromolecule has been spin labelled *in silico*, PELDOR-derived distances can be used to localise individual spin centres such as paramagnetic metal ions or EPR-active substrates of enzymes relative to the model of the macromolecular structure [13,14]. Although the location of metal ions can in most cases be inferred from sequence alignments or crystal structures, in some cases this problem is far from trivial, even if high-resolution structures are available. An example of this is metallonucleases where multiple metal centres can be part of the active site and even movements of the metals during the reaction cycle are proposed [15]. Although crystal structures of these enzymes are available, the location and function of the metal ions are still under debate [15]. Similarly, the localisation of the active site of a novel enzyme (especially non-metallo enzymes) can be difficult, for example, when only a structure of the apo-enyzme is available. The trilateration of spin-labelled or intrinsically EPR-active substrates can in such cases provide a quick and cost-effective solution to locate the active site or a metal-binding site. The general idea of trilateration is the localisation of a certain object in three-dimensional (3D) space by measuring distances between this object and known reference objects [16]. In the context of EPR spectroscopy, spin labels introduced by SDSL would serve as reference objects for the trilateration and the object to be localised would be the intrinsic spin centre, such as a metal ion [14], the spin-labelled or intrinsically EPR-active substrate [13] or additional spin labels [17,18]. Another important application for PELDOR-derived distances is the *in silico* reconstruction of macromolecular complexes from individual domains. Often, structural information is only available for subunits of such complexes and PELDOR can be used to reconstruct the complex based on distance-constrained rigid-body docking or refinement

*Corresponding author. Email: hagelueken@pc.uni-bonn.de

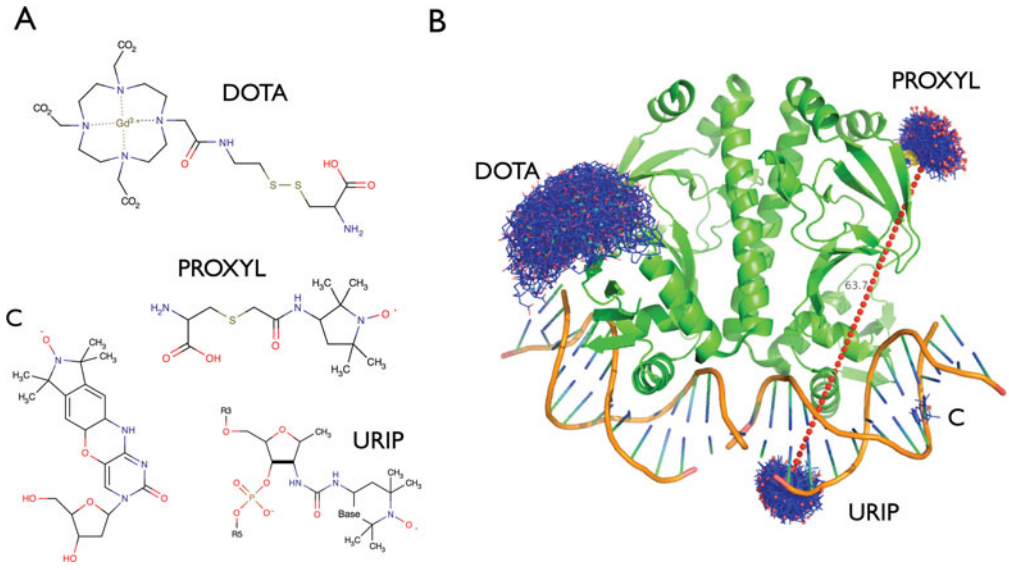


Figure 1. New features implemented in mtsslWizard. (A) Structures of the four newly included spin labels. (B) Examples for the use of the spin labels listed in (A). The dashed red line between the PROXYL and URIP labels represents the average distance between these two ensembles. The old version of the program showed all possible distances as lines, which resulted in screen cluttering.

[19–22]. Results of such experiments can give important insights into the geometry of macromolecular complexes, and thus, serve as a basis for further experiments and structural analysis.

We present here the software package mtsslSuite, which consists of three programs – mtsslWizard, mtsslTrilaterate and mtsslDock. These programs are designed to construct spin-labelled models of macromolecules (mtsslWizard) as well as to use these models for the localisation of spin centres by trilateration (mtsslTrilaterate) or the reconstruction of macromolecular complexes by distance-constrained rigid-body docking (mtsslDock). All three programs can provide scientists with valuable feedback to decide if the available distance constraints support or disagree with a current working model, and if, for example, more labels have to be introduced to get an accurate answer to the problem at hand.

Results and discussion

All programs of the mtsslSuite are designed to work in conjunction with the freely available PyMOL molecular graphics program (www.pymol.org). MtsslTrilaterate and mtsslDock can use the models of spin-labelled macromolecules generated by means of mtsslWizard or other packages as input data. In the following, the usage and the special features of each program will be discussed with the help of examples.

mtsslWizard

The mtsslWizard program allows the user to spin label macromolecules *in silico*. Its graphical user interface (GUI) is fully integrated into PyMOL. Generally, the user can choose the type of spin label from the mtsslWizard menu (supplementary Figure 1) and then attach the chosen label to the selected residue of a biomolecule by clicking on the amino acid/nucleotide it should be positioned at [10]. The program rotates the spin label around its rotatable bonds in order to explore the volume, which is accessible to the label (‘tether-in-a-cone-model’, [23–27]). MtsslWizard does not consider preferred rotamers of the spin label but only excludes conformations that produce clashes with the macromolecular surface or with a label itself, considering a user-defined cut-off value. The advantages and limitations of this approach are thoroughly discussed in a previous publication [10]. In the following, only the changes between versions (1.0) and (1.1) are discussed.

In the first version of MtsslWizard (1.0), only the spin-label MTSSL was included into the program and it was shown that experimental data from a large test data-set (52 distances) can be predicted with an average accuracy of 3 Å, which is on par with other available software packages (MMM, PRONOX). Here, we present an improved version of mtsslWizard (1.1), which contains several new features in comparison with the original version (1.0). Firstly, the algorithm has been optimised, so that the calculation speed has increased and is no longer dependent on the size of the macromolecular model. Secondly, the GUI has been slightly modified in order to make it more comprehensible (supplementary Figure 1). For example, the ‘allowed clashes’ criterion has been merged with the ‘vdW-cut-off’ criterion. The resulting ‘vdW-restraint’ setting can now be switched between two values: ‘tight’ (3.4 Å cut-off, 0 clashes) and ‘loose’ (2.5 Å cut-off, 5 clashes). The ‘loose’ setting allows closer contacts between label and protein, accounting

for flexibility of the system. This setting is useful if a site can be experimentally labelled but cannot be labelled *in silico* using the default 'tight' setting, because the static molecular model is too crowded. As discussed previously [10], the value of 2.5 Å was chosen as a minimum cut-off value because it represents the distance between the donor and acceptor of a short hydrogen bond, whereas the 3.4 Å cut-off represents a short vdW interaction (ignoring the hydrogens). The mtsslWizard asks the user to switch to the 'loose' setting if less than 10 conformations of a label can be found during a run. Importantly, the new release of mtsslWizard includes several additional spin labels that were requested by users (Figure 1(A)): the PROXYL spin label, the gadolinium-based DOTA spin label [28] and two spin labels for nucleic acids, the C-label [29] and a urea-based TEMPO spin label (URIP) [30]. Figure 1(B) shows examples for use of the newly included spin labels. Since mtsslWizard is based on the accessible volume approach, any spin or FRET label with known structure can be easily added as a plug-in to the program, even if no rotamer libraries are available.

Version (1.1) of the program was benchmarked against the original test data-set [10] consisting of 52 MTSSL-derived distances (Figure 2(A)). The results confirm that the program can predict experimental data with an accuracy of 3 Å, similar to MMM or PRONOX (see also [31]). The slight deviations in the prediction of some of the interspin distances between this version of mtsslWizard and the original version are due to the simplified clash criterion (see above). For the newly added spin labels, no large experimental data-sets are available, and we therefore cannot provide a similarly extensive benchmark as for the MTSSL spin label. For the PROXYL label, we used mtsslWizard to predict experimental distance distributions from PROXYL-labelled light-harvesting complex II [12]. The overlay of the experimental distance distributions with the ones from mtsslWizard (Figure 2(B)) shows that the program can predict the PROXYL-based distributions with good accuracy. The DOTA label was tested using a doubly DOTA-labelled transmembrane peptide resulting in good fits of experimental distance distributions [32]. To test the C spin label, a model of a DNA duplex (make-NA server, http://structure.usc.edu/make-na/server.html) was constructed using the sequence of ds-DNA1 (fwd: 5′-GATGCGFGCGCGCGACTGAC-3′, rev: 3′-CTACGCGCGCGCGCTGAFTG-5′) from [33] and the C spin label was attached to the positions marked by 'F' in the sequence using mtsslWizard. The distance between the modelled C labels is 37.6 Å, close to the experimental value of 36.5 Å Figure 2(D). Note that due to the rigidity of the C label, it is only attached to the selected nucleotide by mtsslWizard and not rotated in any way. As a test for the URIP label, a double-stranded polyA:T 15mer was experimentally spin labelled using URIP at positions 3 and 8 of the T strand and PELDOR data were recorded (Figure 2(C)). A model of the labelled DNA strand was constructed using the make-NA server (http://structure.usc.edu/make-na/server.html), PyMOL and mtsslWizard. Figure 2(C) shows that experimental data (mean distance 27.7 Å [DeerAnalysis) and prediction (mean distance 27.0 Å) fit well.

mtsslTrilaterate

The mtsslTrilaterate program can be used to locate spin centres in the structure of biomolecules by means of trilateration, for example, to find metal- or substrate-binding sites (see above).

Once the trilateration program has been launched via the PyMOL menu, its GUI appears as shown in supplementary Figure 2(A). The trilateration process can be prepared by importing the attached spin labels and experimental distances into the program as explained in the software manual (www.pymolwiki.org). The distance data itself can be entered in two different ways: manually or by import of data files. The program can, for example, import DeerAnalysis [34] interspin distance distributions. The program solves the trilateration problem in two steps. In the first step, the coordinates of the unlocalised spin centre are estimated by means of singular value decomposition (SVD). The result of this step is needed as an 'initial guess' for the nonlinear least squares (NLS) optimisation. The goal of NLS algorithms is the minimisation of the commonly used χ^2 merit function,

$$\chi^2(x, y, z) = \sum_l \left[\frac{r_1 - \sqrt{(x - x_l)^2 + (y - y_l)^2 + (z - z_l)^2}}{\sigma_l} \right]^2, \quad (1)$$

where l is the number of the spin label, $(x_l y_l z_l)$ are coordinates of the l spin label, r_l and σ_l are the mean value and the standard deviation of the distance between the l spin label and the unlocalised spin centre. Here, (xyz) are the unknown coordinates of the spin centre. They are incremented in each iteration and the corresponding value of χ^2 is calculated. Then, the program checks whether the obtained χ^2 is decreased by a factor of less than 10^{-3} after the last iteration. If this criterium is satisfied, the problem is considered to be solved. The NLS algorithm uses either the inverse-Hessian method or the Levenberg–Marquardt method [35]. The user can choose the algorithm in the 'Preferences' menu (supplementary Figure 2(B) and (C)). In our hands, both algorithms produce exactly the same results. For each algorithm, a number of parameters can be set in the 'Preferences' menu (supplementary Figure 2(B) and (C)). These settings include the maximum number of iterations, the minimal χ^2 value (below which the problem is considered solved) and the damping parameter lambda (used only for the Levenberg–Marquart algorithm). The values used as default were found to be optimal for the test cases described below. Further description of these parameters can be found elsewhere [35]. The

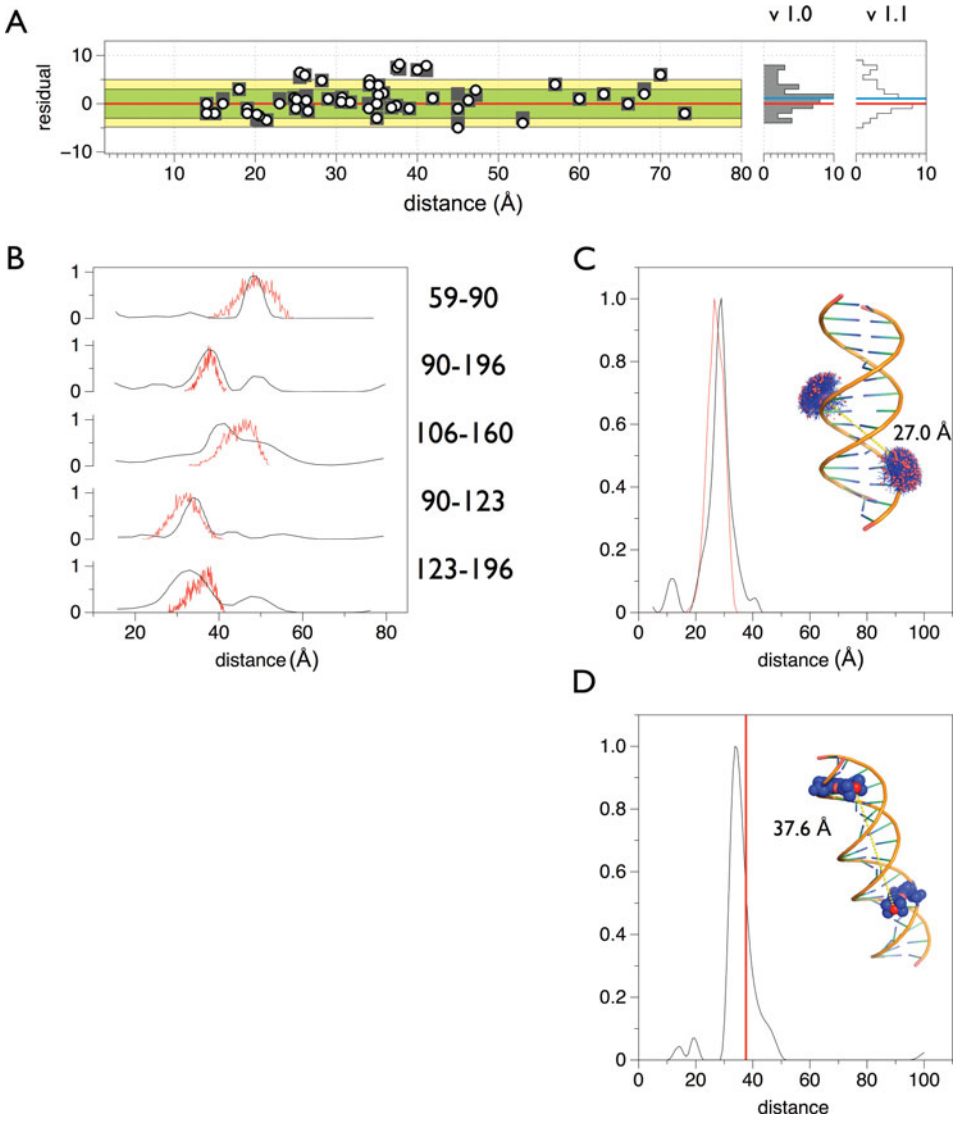


Figure 2. Benchmarking of mtsslWizard. (A) Comparison of version (1.0) of mtsslWizard (grey boxes) with version (1.1) (white circles). Experimental distances that were included in the benchmark data-set of the original mtssWizard paper are plotted against predictions of versions (1.0) and (1.1) of mtsslWizard. The *x*-axis shows the experimental distance and the *y*-axis shows the difference of the experimental value to the prediction. The ideal $y = 0$ line is marked in red, the areas corresponding to different prediction errors are shaded with different colours (green: <3 Å, yellow: ≤5 Å and white: >5 Å). (B) Comparison of experimental distance distributions of the PROXYL-labelled light harvesting complex II (black line, digitised from [12], PDB-ID: 2BHW) with mtsslWizard predictions (red line). (C) Example for the use of the URIP label. The experimental distance distribution is shown in black and the prediction based on the model shown in the inset is shown in red. (D) Example for the use of the C spin label. The experimental distance distribution (black) was constructed using data from [33]. The red vertical line highlights the distance that was calculated by mtsslWizard. The inset shows the DNA (cartoon model) and the modelled spin label (spheres).

solution of the trilateration problem can be interpreted as the most probable coordinates of the unlocalised spin centre and their standard errors. These data are presented as a table in the 'Output' panel (supplementary Figure 2(A)) and can optionally be visualised graphically inside PyMOL (Figures 3 and 4). The calculation statistics are presented in terms of the obtained χ^2 value and the number of NLS iterations used to reach this χ^2. The χ^2 value shows how precisely the trilateration problem was solved for the given set of input data. A lower value of χ^2 corresponds to a lower uncertainty in the coordinates obtained. The factors that affect the χ^2 value are the number of spin labels used to localise the intrinsic spin centre, the precision and the dispersion of the distances evaluated from PELDOR time traces and the accuracy of the model of the spin-labelled biomolecule.

We tested the mtsslTrilaterate program for soybean seed lipoxygenase-1 (SBL1) using a data-set of

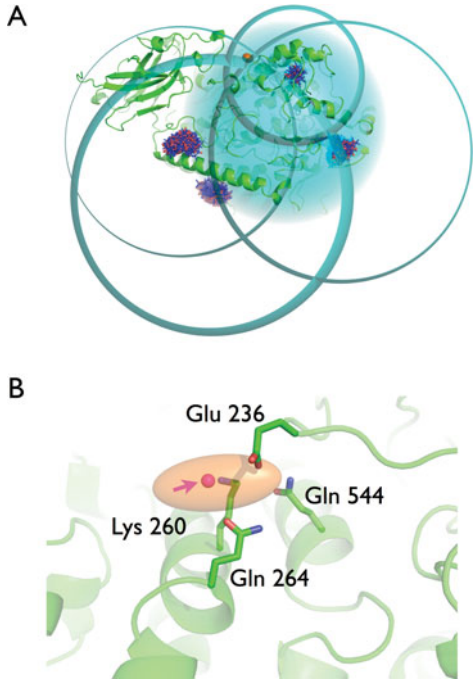


Figure 3. Trilateration of the LOPTC lipid in soybean seed lipoxygenase-1 (SBL1). (A) SBLI is shown as green cartoon model. The MTSSL ensembles that were generated with mtsslWizard are shown as blue and red stick models. The trilateration spheres whose radii correspond to the experimental LOPTC-MTSSL distances are shown in cyan. The trilateration result ('target') is represented as an orange ellipsoid and marks the position of the LOPTC lipid. (B) Close up of the 'target' area. The protein is shown as green cartoon model; selected amino acid residues are shown as sticks. The target is represented by a translucent orange ellipsoid. The magenta spheres inside the ellipsoid are the results of 10 independent attempts of docking LOPTC into SBLI with mtsslDock (marked by arrow).

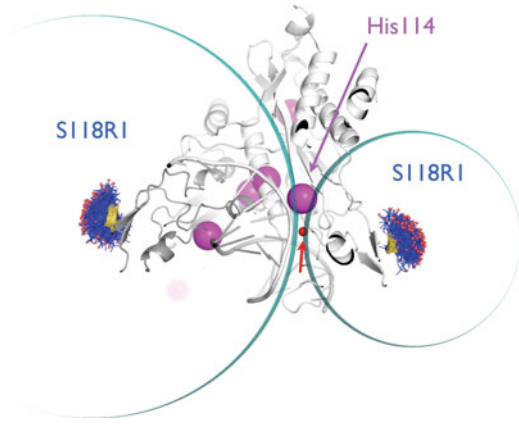


Figure 4. Trilateration of the inhibitory copper-binding site in EcoRI. The EcoRI:DNA complex is shown as a white cartoon model. The MTSSL ensembles that were generated with mtsslWizard are shown as blue and red stick models. The trilateration spheres whose radii correspond to the experimental Cu-MTSSL distances are shown in cyan. All histidine residues in the dimeric structure are highlighted as purple spheres. His114 that was identified as the Cu^{2+} binding site in [14] is marked and is in the closest proximity to both trilateration spheres. The red spheres (marked by a arrow) represent the result of 10 independent attempts of docking the Cu^{2+} ion into the complex structure with mtsslDock.

PELDOR-derived distances published by Gaffney *et al.* [13]. The data-set consists of five distances measured between the TEMPO-labelled lipid (LOPTC) and five spin labels that were attached to lipoxygenase by SDSL. In the original paper, these data were used to localise the polar end of the LOPTC on the molecular surface of lipoxygenase by means of trilateration using multi-dimensional scaling and Procrustes analysis that are included in the MATLAB (http://www.mathworks.de) statistical toolbox. The mtsslTrilaterate program was run with the input parameters listed in Table 1. The coordinates of the spin labels were obtained by using mtsslWizard and the crystal structure of SBL1 (PDB:1YGE, [36]). The mean values and standard deviations of the interspin distances between LOPTC and five spin labels were taken from [13]. Calculations were performed with the default settings of mtsslTrilaterate (see supplementary Figure 2(B) and (C)). The results are shown in Figure 3 and Table 1. As can be seen from Figure 3, the program places the LOPTC spin centre next to helices 2 and 11 of SBL1. In this position, LOPTC is surrounded by amino acids Glu236, Lys260, Gln264 and Gln544 (Figure 3(B)). Within the error of calculations, the trilateration result coincides with those published in the original paper (Table 1). A slight deviation between these two solutions may be related to the usage of different models of the spin-labelled protein (either derived from mtsslWizard or PRONOX) and different calculation procedures. To estimate the influence of differences in the models on the solution, the mtsslTrilaterate program was run with the coordinates of the spin labels taken from [13]. It was found that the coordinates of the LOPTC spin derived in this way fit with high precision to the published coordinates. This fact indicates that the small deviations between the solutions are related to the differences in the models of spin-labelled protein. With mtsslTrilaterate, this localisation of LOPTC was done within 10 min, including the *in silico* spin labelling of the labelled positions with mtsslWizard.

As a second test case, we took an experimental data-set that was used to localise an inhibitory copper-binding site in the restriction endonuclease EcoRI [14]. Position Ser180 in both monomers of the dimeric protein (PDB-ID: 1CKQ) was labelled with mtsslWizard and the experimental distances between the spin label and the copper site to be localised were used as input for mtsslTrilaterate. Although these input data are not sufficient for the program to solve the trilateration with a unique solution (at least four distances are needed for this), the visualisation of the trilateration spheres in PyMOL (Figure 4) shows that the

Table 1. Calculated coordinates of spin centres in the molecular coordinate system of soybean seed lipoxygenase-1 (PDB:1YGE) and PELDOR-derived distances between these centres and LOPTC.

Spin centres	Program used	Coordinates (PDB:1YGE) (Å)			PELDOR-derived distances (Å)
		x	*y*	*z*	
$F270R_1$	mtsslWizard	16.66	61.53	0.04	23.0 ± 6.2[a]
	PRONOX[a]	16.26[a]	62.24[a]	−0.63[a]	
$L480R_1$	mtsslWizard	−6.80	44.17	1.01	43.0 ± 2.3[a]
	PRONOX[a]	−8.09[a]	44.49[a]	−0.48[a]	
$A569R_1$	mtsslWizard	11.18	45.17	33.08	38.0 ± 5.9[a]
	PRONOX[a]	10.63[a]	42.72[a]	34.13[a]	
$A619R_1$	mtsslWizard	26.15	14.97	9.29	47.0 ± 2.3[a]
	PRONOX[a]	26.66[a]	15.88[a]	8.96[a]	
$F782R_1$	mtsslWizard	42.22	28.14	0.08	37.5 ± 5.9[a]
	PRONOX[a]	41.56[a]	28.93[a]	−0.30[a]	
LOPTC	mtsslTrilaterate[b]	32.4 ± 2.0	61.6 ± 1.6	9.6 ± 4.6	
	Gaffney *et al*.[c]	30.4 ± 2.2[a]	62.2 ± 2.3[a]	12.2 ± 6.0[a]	
	mtsslTrilaterate[d]	30.0 ± 2.2[b]	62.6 ± 2.3[b]	11.6 ± 4.2[b]	
	mtsslDock[e]	32.6[c]	62.0[c]	9.1[c]	

[a]Taken from [13]. The standard deviations of the interspin distances were estimated from the plots of distance distribution assuming that the distributions have a Gaussian profile.
[b]MtsslTrilaterate was run using mtsslWizard-generated labels as input.
[c]Results from the original publication [13].
[d]MtsslTrilaterate was run for the input data taken from [13].
[e]MtsslDock result for the LOPTC position using the mtsslWizard-generated labels as input.

point of closest approach of the two spheres coincides with His114, which was also identified as the Cu^{2+} binding site in the original paper.

mtsslDock

The mtsslDock program enables the user to dock two macromolecules based on a set of experimental PELDOR distances. The program requires models of the spin-labelled macromolecules and the experimental distance data between them as input. The structural models can be spin labelled with mtsslWizard or other available software packages.

After launching mtsslDock from within PyMOL, its GUI appears as shown in supplementary Figure 3. The docking process is then prepared by importing the spin-labelled docking partners into the program as explained in the software manual (www.pymolwiki.org). In the following, the algorithm and the adjustable settings are explained: mtssldock uses a mixed genetic and evolutionary algorithm [37] to find docking models between two macromolecules (e.g. proteins A and B) that agree to the experimental distance data. The principle of genetic and evolutionary algorithms is based on natural evolution where inheritance, mutation and genetic crossover between generations lead to a population with increased fitness for survival in the particular environment. Such algorithms have been used for a wide array of problems, including protein–protein docking based on surface complementarity [38] or molecular replacement in crystallographic phasing [39]. The algorithm works as follows: protein A is held fixed in space, while in the first step of the calculation, mtsslDock generates a population of 400 random orientations (each orientation is called a chromosome) of protein B. A chromosome is a set of six numbers ('genes'), which determine the translation and rotation of the protein in 3D space: $x, y, z, \alpha, \beta, \gamma$ (see also supplementary Figure 4). The root-mean-square difference (RMSD) between the experimental distances and distances derived from the current docking model as well as the χ^2 value (as specified in Equation (1), see above) are then determined. The χ^2 value is used as a surrogate for the fitness of the particular chromosome – here, a low χ^2 value corresponds to a higher fitness than a high χ^2 value. In the next step, the 5% least-fit chromosomes of the population are discarded and replaced by 'offspring' of the surviving chromosomes (supplementary Figure 4). The offspring is generated by genetic operations: random mutagenesis, small creep mutagenesis, exchange crossover or single-point crossover (e.g. [38], supplementary Figure 4), forming a new generation of the population. This process is repeated 1000 times before a 'packing-function' adds a high fitness penalty to chromosomes that lead to protein B clashing into protein A (only C_α atoms are considered; a distance lower than 3.4 Å between two C_α atoms is counted as a clash). The 5% fittest trial orientations are then used as starting orientations for a second evolution with 50 generations in which only small creep mutations (supplementary Figure 4) are introduced to the fittest chromosome to replace the 95% unfittest chromosomes in each step. Here, the fitness is determined by scoring χ^2 as well as clashes between both proteins. This effectively resembles a least-squares rigid-body refinement. In some cases,

conformational changes of the docking partners between the bound- and unbound state might lead to 'false negative' docking results due to the clash criterion. To account for this, mtsslDock does not simply discard clashing solutions but displays them marked as 'clashing' inside the GUI and within PyMOL, so that ultimately the user decides whether a solution should be discarded or not.

By default, the program calculates 10 independent populations and exports the fittest solution of each population to PyMOL to be investigated by the user. A docking calculation takes between 0.5 and 5 min for a single population on a standard laptop, depending on the settings that were used and the size of the molecules (due to the clash scoring). We recommended to do the calculation with at least 10 independent populations to get an impression of how statistically significant the produced solutions are.

The robustness of the algorithm was analysed using the crystal structure of the protein–protein complex rubredoxin:rubredoxin-reductase (PDB-ID: 2v3b, [40]). The rubredoxin reductase was kept static (protein A), while rubredoxin was allowed to be moved by mtsslDock (protein B). Both proteins were spin labelled at 12 positions with mtsslWizard (supplementary Figure 5), and 4–12 of the 144 possible distances were randomly selected as constraints. Each label was only allowed to contribute to one distance constraint. Because of the statistic nature of the docking algorithm, the results of this analysis have to be regarded in terms of the ability of the algorithm to reliably find a correct docking solution with a given number of constraints. For the purpose of our benchmark, we arbitrarily define a correct solution as one that has an RMSD (all atoms) to the experimental position of protein B of less than 2.5 Å. Figure 5(A) (top) shows that in each docking run of our benchmark, the algorithm found potential solutions that fit to the 'experimental' distances with RMSD values <1 Å. However, only when six or more constraints were used, was the correct solution among the docked structures (Figure 5(A), bottom). With an increasing number of constraints, the fraction of wrong solutions decreased, simply because it becomes less likely for the algorithm to find an additional solution that satisfies all distance constraints (Figure 5(A), bottom). Figure 5 (B) and (C) shows the docking solutions when 7 or 10 constraints were used, respectively. Whereas all solutions were correct when 10 constraints were used (Figure 5(C)), the 10 solutions derived from the 7 constraints form 2 clusters (Figure 5(B), red and green), one of which is located at the correct position. For both the clusters, all solutions show a similar fit to the 'experimental data' with RMSD <1 Å (Figure 5(A), top). Importantly, this is much lower than the uncertainty that is on average introduced by *in silico* spin labelling ($\sim$3 Å) [10,31] and experimental errors. Figure 5(D) shows that interestingly, the 'worst' solution of the correct cluster has a χ^2 value that is higher than that of the incorrect solutions. Thus, in reality, right and wrong solutions cannot simply be distinguished by ranking the solutions. In our benchmark case, it would be straightforward to identify the correct cluster of protein B, since the wrong solutions do not make any contact with protein A. However, in reality and especially in cases, where, for example, a domain of the protein complex is missing from the structure, more constraints would have to be added to get a definite answer.

Proteins often form functional dimers and sometimes only the monomeric structure or competing dimer models can be derived from crystallography. PELDOR can in such cases be used to determine the native dimer structure [22]. We thus added an option to mtsslDock that only C2 symmetric docking models are constructed and evaluated during the docking run. In theory, less constraints (and therefore experimental time) should be needed in such cases, because the system has less degrees of freedom. We analysed this function of mtsslDock with the same benchmark, but used the structure of the dimeric Cap protein (Figure 1(A), PDB-ID: 1O3Q) as a test case. This time, six labels were attached to each monomer (at corresponding positions) of the protein (supplementary Figure 5) and six, five or four distances between monomers were randomly chosen as constraints. Figure 5(A) (blue dots) shows, that as expected, correct solutions can be found with down to four constraints. More symmetry restraints will be added in future releases of the software.

In summary, our algorithm can reliably dock two structures when enough distance constraints are available (six or more for unsymmetric structures, four or more for C2 symmetric structures). It cannot distinguish a correct from an incorrect solution when distances derived from both solutions show a similarly good fit to the experimental distances. This situation can arise when not enough constraints are present, when the algorithm gets trapped in local minima or from specific placements of labels. For example, when three labels were placed on protein A, they would form a mirror plane for any number of labels on protein B, and the algorithm would not be able to distinguish the solution from its mirror image.

We also tested mtsslDock with three published sets of experimental interspin distances. The first example is the complex between the cytoplasmic domain of erythrocyte band 3 and ankyrin-R repeats 13–24 [20]. This complex has been investigated by various biochemical and biophysical techniques, including a set of 20 PELDOR measurements. In the original paper, the PELDOR-derived distances were used as distance constraints for docking analysis with RosettaDock [41], resulting in a model of the protein complex that is in agreement with other biochemical and biophysical data [20]. To dock the complex with mtsslDock, we labelled the experimentally used sites in both proteins with mtsslWizard, imported the mean coordinates of the labels into mtsslDock and entered the 20 experimental distances and standard deviations [20]. The result of the

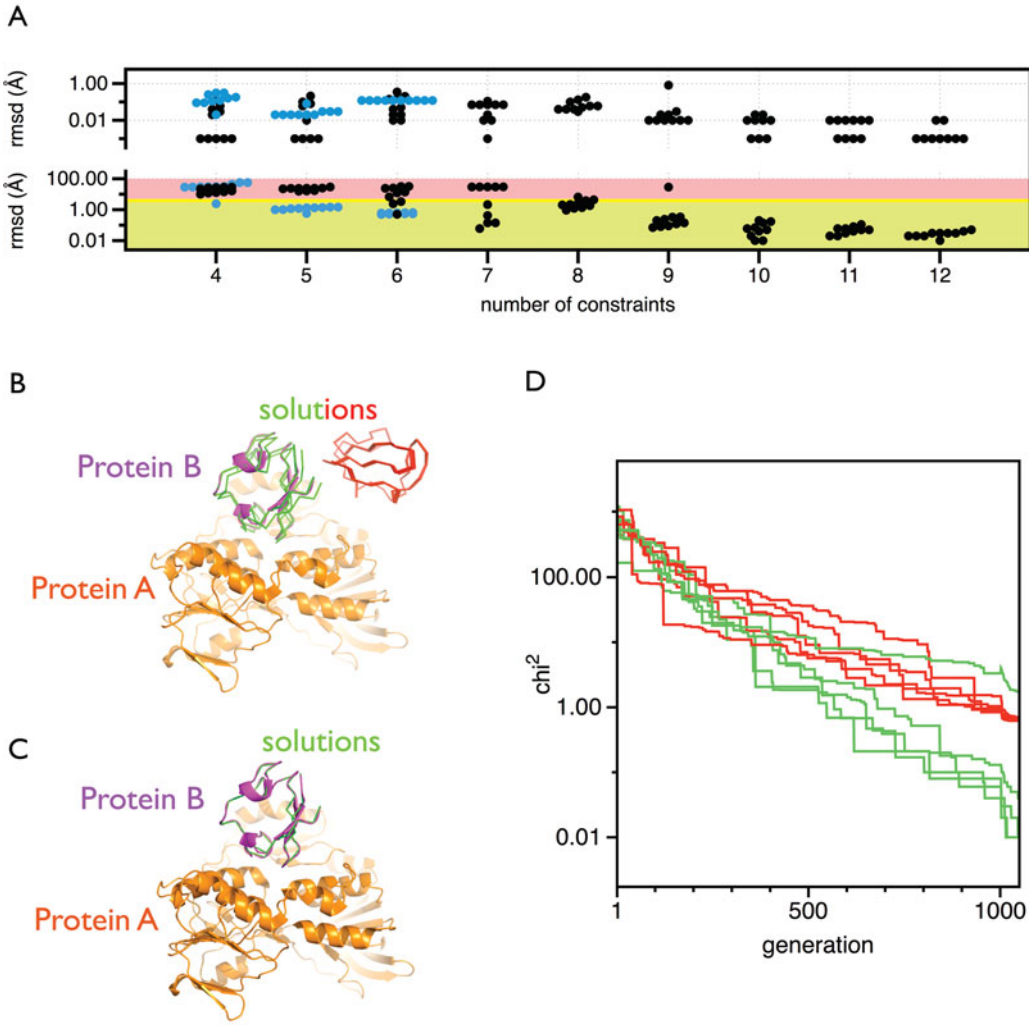


Figure 5. Analysis of the mtsslDock algorithm. (A) Benchmark results. Top: Each black or blue dot represents the result of an independent docking run that was performed with the number of constraints indicated on the *x*-axis. The *y*-axis (logarithmic scale) represents the RMSD between the 'experimental' distances and the distances retrieved from the docking result. The black dots represent the docking runs performed with rubredoxin:rubredoxin-reductase, and the blue dots, the docking runs performed with Cap (see main text). Bottom: The graph is analogue to the top panel but shows the RMSD difference between the docking result and the experimentally found position of the docked molecule. The area with green shading represents RMSD values <2.5 Å, yellow: ≤5 Å, red: >5 Å. (B) Docking results (red, green) for rubredoxin:rubredoxin-reductase (magenta:orange) with seven constraints. (C) Same as B, but 10 constraints were used. (D) Fitness traces of the docking runs for the solutions shown in (B). The traces are coloured according to the solutions in (B).

docking calculation (10 docking runs were used, calculation time: 30 min) is shown in Figure 6 and resembles the result that was obtained in the original publication (supplementary Figure 5). Interestingly, the program produces two clusters of possible solutions, which have similar fitness scores (Figure 6, supplementary Figure 5). Both of these clusters are contained in the ensemble of the published solutions. Since the number of constraints in this case greatly exceeds six (see above and Figure 5(A)), it seems surprising that our algorithm does not find a single solution. There are several possible reasons for this: (1) some of the distances might be incompatible with each other due to *in silico* spin-labelling uncertainties and/or experimental uncertainties, (2) the algorithm might get trapped in local minima or (3) some of the constraints are not linearly independent from each other. The latter reason could arise from the fact that in this case most labels were (understandably) repeatedly used for distance measurements. More linearly independent constraints would have to be introduced to resolve this ambiguity. In the original publication, the authors did this by introducing constraints from orthogonal methods, such as solvent accessibility and cross-linking data.

The mtsslDock approach can also be used to solve the trilateration problem (see above). For this purpose, we used the SBL-1 example (see mtsslTrilaterate above) and imported the mean coordinates of the five spin labels, the

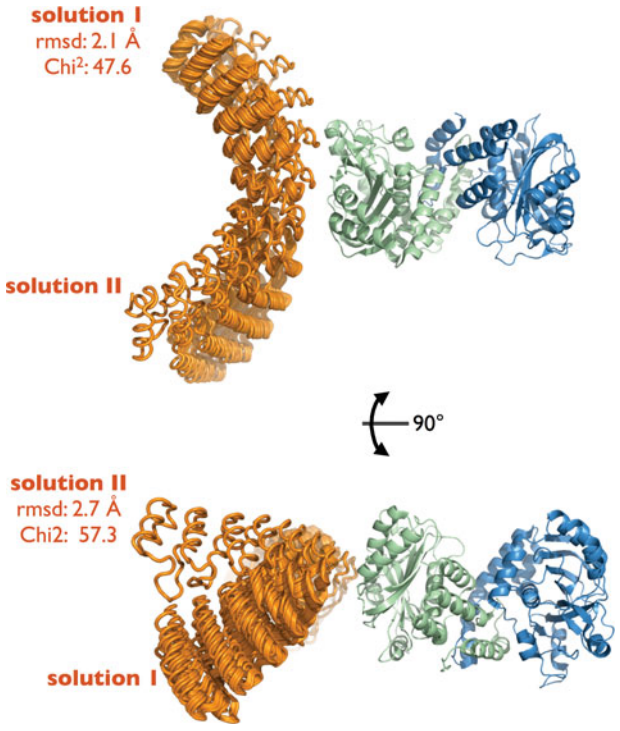


Figure 6. Docking of the cytoplasmic domain of erythrocyte band 3, and ankyrin-R repeats 13–24 [20]. The docking solutions for the ankyrin structure based on 20 PELDOR distances are shown as orange loop models. The erythrocyte band 3 dimer is shown as cartoon model (green and blue). MtsslDock finds two solutions for the problem (solution I and solution II). In the bottom panel, the structure is rotated by 90°.

coordinates of an atom that was used as a surrogate for the position of the LOPTC lipid (termed 'pseudoatom' below) and the five experimental interspin distances into mtsslDock. The docking was performed using the default settings. Figure 3(B) (purple spheres) shows that the program places the pseudoatom into the centre of the 'target' ellipsoid that was generated by mtsslTrilaterate. Similarly mtsslDock was also used to compute the position of the copper-binding site in EcoRI (see above, [14]). The result of this calculation lies exactly between the two trilateration spheres that were output by mtsslTrilaterate and close to His114, which was identified as the copper-binding site in the original publication (Figure 4, red spheres). Note that the program would output a range of differing solutions spread on an intersection circle if the two triangulation spheres would intersect each other. Thus, this result has to be considered as a special case. These two examples represent a cross-check of both algorithms. The use of mtsslTrilaterate is, however, recommended in such cases since it is faster, calculates statistics and provides more informative output for the trilateration problem.

Conclusion

The mtsslSuite currently contains a set of three applications that allow *in silico* spin labelling of macromolecules, localising spin centres with respect to reference spin centres or docking macromolecules based on experimental spin-pair distance constraints.

The programs can import spin-labelled locations that were generated with mtsslWizard but do at the same time offer users the opportunity to use input data from other sources (e.g. MMM or PRONOX). We have shown that all the three programs achieve results that are comparable in precision to those obtained by means of other computational strategies. The source code of all the three programs is freely available at (www.pymolwiki.org).

Acknowledgements

We would like to thank Jason Vertrees and Thomas Holder for help concerning PyMOL programming and Damien Farrell for help with the tkinterTable class for mtsslDock. This work has been supported by the SFB813 projects C8 and A6 of the Deutsche Forschungsgemeinschaft (DFG) and the Wellcome Trust (091825/Z/10/Z).

Note added in proof

While this paper went into production, the K1 spin label [42] was added to mtsslWizard upon user request. Supplementary Figure 7 shows a comparison between predicted and experimental distance distributions derived from this spin label.

References

[1] G. Jeschke, Annu. Rev. Phys. Chem. **63**, 419 (2012).
[2] G.W. Reginsson and O. Schiemann, Biochem. J. **434**, 353 (2011).
[3] O. Schiemann and T.F. Prisner, Q. Rev. Biophys. **40**, 1 (2007).
[4] A. Milov, K. Salikohov, and M. Shirov, Fizika. Tverdogo. Tela. **23**, 975 (1981).
[5] G. Jeschke, A. Bender, and H. Paulsen, J. Magn. Reson. **169**, 1–12 (2004).
[6] C. Altenbach, T. Marti, H.G. Khorana, and W.L. Hubbell, Science **248**, 1088 (1990).
[7] J.P. Klare and H.-J. Steinhoff, Photosyn. Res. **102**, 377 (2009).
[8] S.A. Shelke and S.T. Sigurdsson, Eur. J. Org. Chem. **2012**, 2291 (2012).
[9] R. Ward and O. Schiemann, *Structural Information from Oligonucleotides* (Springer, Berlin, 2012).
[10] G. Hagelueken, R. Ward, J.H. Naismith, and O. Schiemann, Appl. Magn. Reson. **42**, 377 (2012).
[11] M.M. Hatmal, Y. Li, B.G. Hegde, P.B. Hegde, C.C. Jao, R. Langen, and I.S. Haworth, Biopolymers **97**, 35 (2011).
[12] Y. Polyhach, E. Bordignon, and G. Jeschke, Phys. Chem. Chem. Phys. **13**, 2356 (2011).
[13] B.J. Gaffney, M.D. Bradshaw, S.D. Frausto, F. Wu, J.H. Freed, and P. Borbat, Biophys. J. **103**, 2134 (2012).
[14] Z. Yang, M.R. Kurpiewski, M. Ji, J.E. Townsend, P. Mehta, L. Jen-Jacobson, and S. Saxena, Proc. Nat. Acad. Sci. U.S.A. **109**, E993–1000 (2012).
[15] C.M. Dupureur, Curr. Opin. Chem. Biol. **12**, 250 (2008).
[16] A. Muschielok, J. Andrecka, A. Jawhari, F. Brückner, P. Cramer, and J. Michaelis, Nat. Methods **5**, 965 (2008).

[17] P.P. Borbat, H.S. Mchaourab, and J.H. Freed, J. Am. Chem. Soc. **124**, 5304 (2002).
[18] G. Jeschke, A. Bender, T. Schweikardt, G. Panek, H. Decker, and H. Paulsen, J. Biol. Chem. **280**, 18623–18630 (2005).
[19] D. Hilger, Y. Polyhach, E. Padan, H. Jung, and G. Jeschke, Biophys. J. **93**, 3675 (2007).
[20] S. Kim, S. Brandon, Z. Zhou, C.E. Cobb, S.J. Edwards, C.W. Moth, C.S. Parry, J.A. Smith, T.P. Lybrand, E.J. Hustedt, and A.H. Beth, J. Biol. Chem. **286**, 20746 (2011).
[21] R. Ward, M. Zoltner, L. Beer, H. El-Mkami, I.R. Henderson, T. Palmer, and D.G. Norman, Structure **17**, 1187 (2009).
[22] M. Zoltner, D.G. Norman, P.K. Fyfe, H. El-Mkami, T. Palmer, and W.N. Hunter, Structure **21**, 595–603 (2013).
[23] N. Alexander, M. Bortolus, A. Al-Mestarihi, H. Mchaourab, and J. Meiler, Structure **16**, 181 (2008).
[24] Z. Guo, D. Cascio, K. Hideg, and W. Hubbell, Protein Sci. **17**, 228 (2008).
[25] S.J. Hirst, N. Alexander, H.S. Mchaourab, and J. Meiler, J. Struct. Biol. **173**, 506 (2011).
[26] E. Hustedt, R. Stein, L. Sethaphong, S. Brandon, Z. Zhou, and S. DeSensi, Biophys. J. **90**, 340 (2006).
[27] Z. Zhou, S.C. DeSensi, R.A. Stein, S. Brandon, M. Dixit, E.J. McArdle, E.M. Warren, H.K. Kroh, L. Song, C.E. Cobb, E.J. Hustedt, and A.H. Beth, Biochemistry. **44**, 15115 (2005).
[28] Y. Song, T.J. Meade, A.V. Astashkin, and E.L. Klein, J. Magn. Reson. **210**, 59 (2011).
[29] P. Cekan, A.L. Smith, N. Barhate, B.H. Robinson, and S.T. Sigurdsson, Nucleic Acids Res. **36**, 5946 (2008).
[30] T.E. Edwards, T.M. Okonogi, and B.H. Robinson, J. Am. Chem. Soc. **123**, 1527 (2001).
[31] G. Jeschke, Prog. Nucl. Magn. Reson. Spectrosc. **72**, 42–60 (2013).
[32] D. Goldfarb (private communication).
[33] O. Schiemann, P. Cekan, D. Margraf, T.F. Prisner, and S.T. Sigurdsson, Angew. Chem. Int. Edit. **48**, 3292 (2009).
[34] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. Timmel, D. Hilger, and H. Jung, Appl. Magn. Reson. **30**, 473 (2006).
[35] B.P. Flannery, W.H. Press, and S.A. Teukolsky, *Numerical Recipes in C*. (Cambridge University Press, New York, 1992).
[36] W. Minor, J. Steczko, O.B. Stec, and Z. Otwinowski, Biochemistry **35**, 10687 (1996).
[37] M. Srinivas and L.M. Patnaik, Computer **27**, 17 (1994).
[38] E.J. Gardiner, P. Willett, and P.J. Artymiuk, Proteins **44**, 44 (2001).
[39] C.R. Kissinger, D.K. Gehlhaar, and D.B. Fogel, Acta Crystallogr. D. Biol. Crystallogr. **55**, 484 (1999).
[40] G. Hagelueken, L. Wiehlmann, T.M. Adams, H. Kolmar, D.W. Heinz, B. Tümmler, and W.-D. Schubert, Proc. Nat. Acad. Sci. U.S.A. **104**, 12276 (2007).
[41] S. Lyskov and J.J. Gray, Nucleic Acids Res. **36**, W233 (2008).
[42] M.R. Fleissner, E.M. Brustad, T. Kálái, C. Altenbach, D. Cascio, F.B. Peters, K. Hideg, S. Peuker, P.G. Schultz, and W.L. Hubbell, Proc. Nat. Acad. Sci. U.S.A. **106**, 21637–21642 (2009).

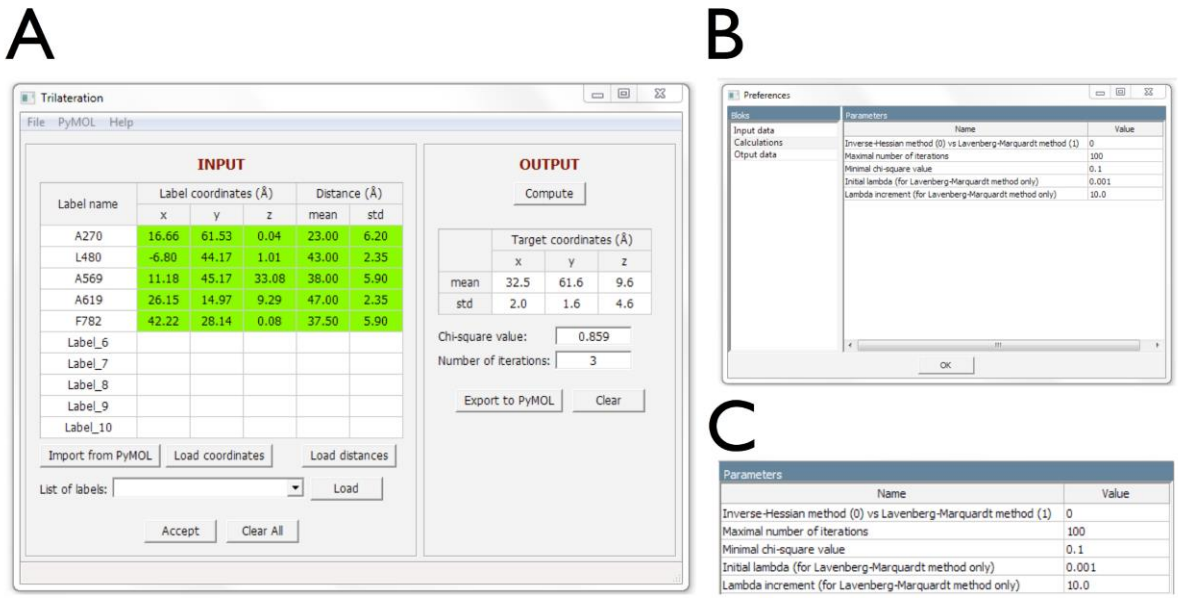
mtsslSuite: *In silico* spin labelling, trilateration and distance-constrained rigid body docking in PyMOL

Gregor Hagelueken[a,*], Dinar Abdullin[a], Richard Ward[b], Olav Schiemann[a,b]

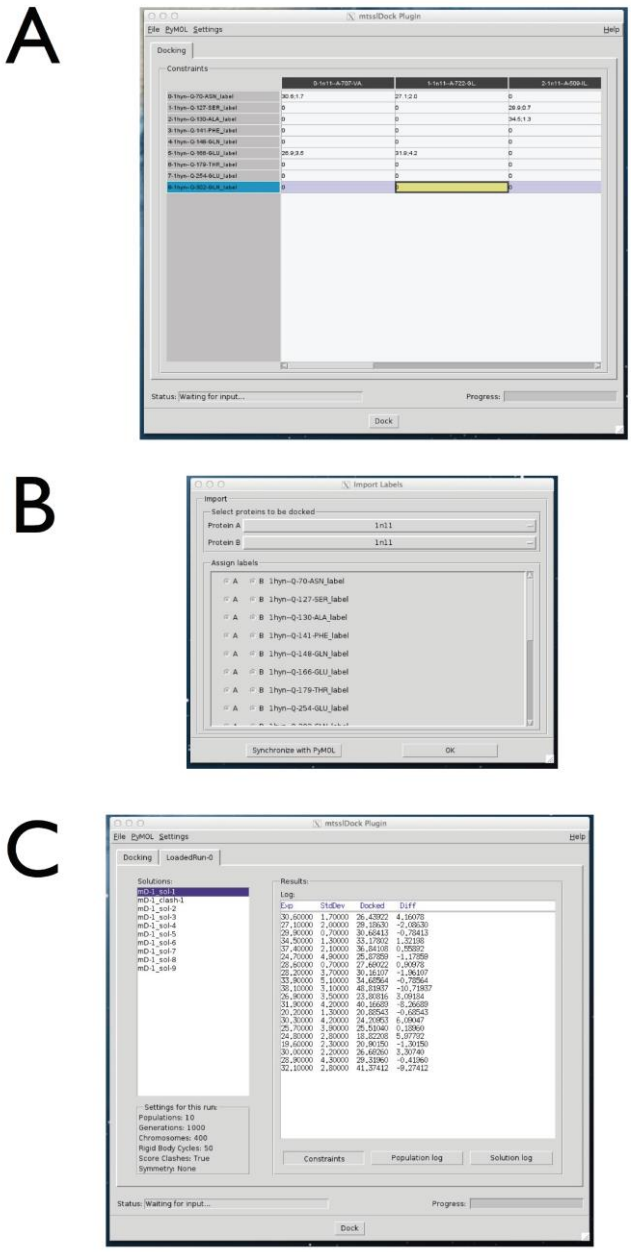
[a]*Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany.*
[b]*School of Physics, University of St Andrews, North Haugh, KY16 9ST St Andrews, UK.*

Supporting Material

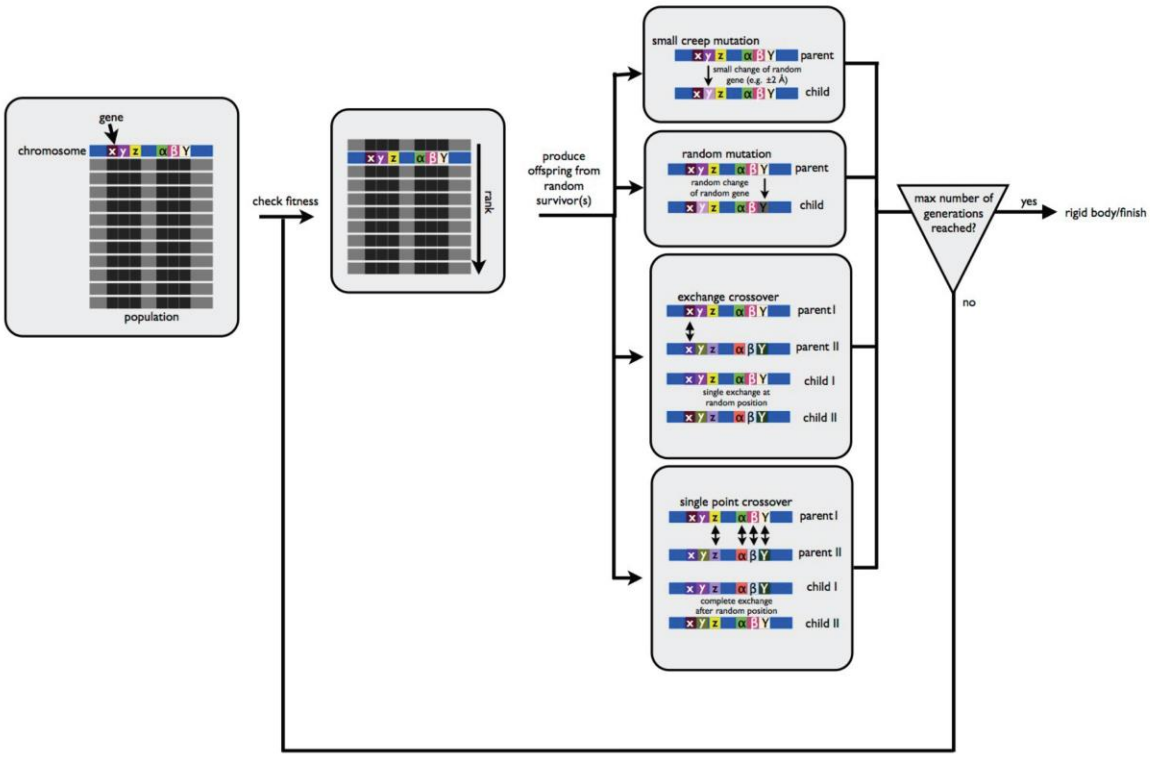
Supplementary Figure 1. The improved mtsslWizard GUI. The inset shows the menu for selecting the new spin labels.



Supplementary Figure 2. The mtsslTrilaterate GUI. A) The main GUI of the program. The 'File' menu on the top can be used to load or save sessions or to open the "Preferences" menu (detailed below). The 'PyMOL' menu offers the possibility to load a coordinate file into PyMOL. The main table (green) in the 'Input' panel shows all currently available data, such as the identifiers and coordinates of spin labels as well as the mean distances and standard deviations of the experimental distance distributions that were assigned to a particular label. These data can be manually typed into the table, imported from PyMOL, loaded from a coordinate file or loaded from a distance distribution file by using the respective buttons on the bottom of the 'Input panel'. The "Output" panel includes the compute button, which is used to start the calculation and a table which displays the results and statistical parameters that are explained in the main text. Clicking the "Export to PyMOL" button generates a graphical representation of the results inside PyMOL (see Figure 4). B) The preferences menu with adjustable parameters for the calculation, which are detailed in the text. C) Blowup of B) showing the default calculation parameters.

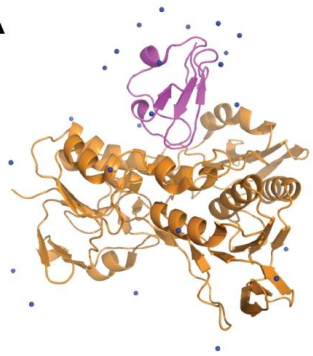


Supplementary Figure 3. The mtsslDock GUI. A) The docking tab harbours the table for experimental distances between spin labels (rows and columns). B) The import tab allows to import spin labels from PyMOL and to assign the imported labels to protein A or B for the docking procedure. C) Docking results page.



Supplementary Figure 4. Schematic of the genetic algorithm used by mtsslDock.

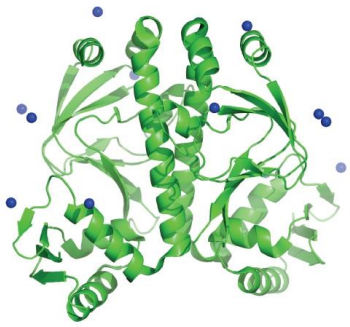
A



Rdx (protein B): 35, 16, 50, 20, 12, 32, 53, 25, 31, 51, 1,29

RdxR (protein A): 340, 65, 311, 218, 178, 90, 249, 119, 306, 97, 31, 140

B



Cap (protein A/B): 22, 44, 158, 17, 37, 152

Supplementary Figure 5. Label positions for the docking benchmark. A) The labelled positions are indicated by blue spheres that correspond to the average position of the label ensemble. The residues that were labeled in each protein are listed on the right. B) The labelled positions are indicated by blue spheres that correspond to the average position of the label ensemble. The residues that were labeled in each protein are listed on the right.

[P3] mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL

Reprinted with permission from

Gregor Hagelueken, Dinar Abdullin, Olav Schiemann, "mtsslSuite: Probing biomolecular conformation by spin labeling studies", *Meth. Enzym.* **2015**, *563*, 595-622.
DOI: 10.1016/bs.mie.2015.06.006

Own contribution to the manuscript:

- Development and testing of the program mtsslTrilaterate.
- Development of the graphical tools for the "Distance Map" mode of the program mtsslWizard.

CHAPTER TWENTY-TWO

mtsslSuite: Probing Biomolecular Conformation by Spin-Labeling Studies

Gregor Hagelueken[1], **Dinar Abdullin, Olav Schiemann**
Institute for Physical and Theoretical Chemistry, University of Bonn, Bonn, Germany
[1]Corresponding author: e-mail address: hagelueken@pc.uni-bonn.de

Contents

Abstract

EPR long-range distance measurements on spin-labeled macromolecules have recently become a popular tool in structural biology. The method can be used to obtain coarse-grained structures of biomolecules, to track conformational changes and dynamics, to dock macromolecular complexes, or to localize spin centers within macromolecules using trilateration. Because the conformation of the spin label is usually unknown, it is often necessary to construct conformational models of the spin label on the macromolecules for data interpretation. For this purpose, so-called *in silico* spin-labeling approaches have been developed. In this chapter, a comprehensive summary of the mtsslSuite is provided, one of the *in silico* spin-labeling software packages. The package currently contains three programs: mtsslWizard, mtsslDock, and mtsslTrilaterate. Worked examples for the usage of all three programs during the planning- and interpretation stages of the EPR experiment are given.

Methods in Enzymology, Volume 563
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2015.06.006

1. INTRODUCTION

The focus of structural biology is shifting toward more and more complex systems. As a consequence, insights that can be gained from the traditional techniques such as X-ray crystallography or NMR are sometimes not sufficient to fully understand a particular system of interest. Important examples are membrane transporters, which are notoriously difficult to crystallize and often too big to be efficiently investigated by NMR. Even if the structure of a transporter can be solved, it represents only one conformational or functional state, which is in most cases not enough to unravel the transport mechanism. Also, crystal structures are usually produced using detergent solubilized membrane proteins (Sonoda et al., 2010). Without doubt, these *in surfo* structures (Caffrey, 2003) provide very important information, but it is also known that the lipid environment of the membrane can be an integral part of the transporters structure and function (Jensen & Mouritsen, 2004). For these reasons, it is important to study a particular system using approaches that provide complementary information (e.g., crystalline state vs. solution/lipid state). Recently, the combination of X-ray crystallography and EPR long-range distance measurements has become very popular, especially in the membrane transporter field (Endeward, Butterwick, MacKinnon, & Prisner, 2009; Georgieva, Borbat, Ginter, Freed, & Boudker, 2013; Hagelueken et al., 2009; Hilger et al., 2005; Joseph, Korkhov, Yulikov, & Jeschke, 2013; Pliotas et al., 2012; Smirnova et al., 2007; Ward et al., 2014; Zou, Bortolus, & Mchaourab, 2009). While crystallography can provide high-resolution information for the complete protein scaffold, EPR-based long-range distance measurements may be used to determine conformational changes of the protein in response to substrate binding or transport events.

Importantly, these experiments can be conducted in different lipid environments such as lipid nanodiscs (Bayburt, Grinkova, & Sligar, 2002), bicelles (Sanders & Prosser, 1998), or liposomes (Sessa & Weissmann, 1968) (practical examples can be found in: Ward et al., 2014; Joseph et al., 2013; Endeward et al., 2009; Gordon-Grossman, Zimmermann, Wolf, Shai, & Goldfarb, 2012). If necessary, the lipid composition of these membrane mimics can be adjusted to mirror the *in vivo* environment. In recent years, EPR distance measurements have even been conducted in living cells (Krstić et al., 2011; Schmidt, Borbas, Drescher, & Summerer, 2014).

Most biomacromolecules are diamagnetic and therefore "invisible" for EPR. Because of this, they have to be spin labeled (Altenbach, Marti, Khorana, & Hubbell, 1990). Different kinds of spin labels exist for different applications and macromolecules. For proteins, the labels are commonly attached to cysteine residues via disulfide linkages, while modified bases are used for nucleic acids such as DNAs or RNAs (Hubbell, López, Altenbach, & Yang, 2013; Klare & Steinhoff, 2009; Reginsson & Schiemann, 2011; Fig. 1). If two or more spin labels are attached to a biomolecule, the dipolar coupling between the two unpaired electrons can be measured with various EPR techniques (Schiemann & Prisner, 2007), for example, pulsed electron–electron double resonance (PELDOR or DEER) (Jeschke, 2012; Milov, Salikohov, & Shirov, 1981). The frequency of the

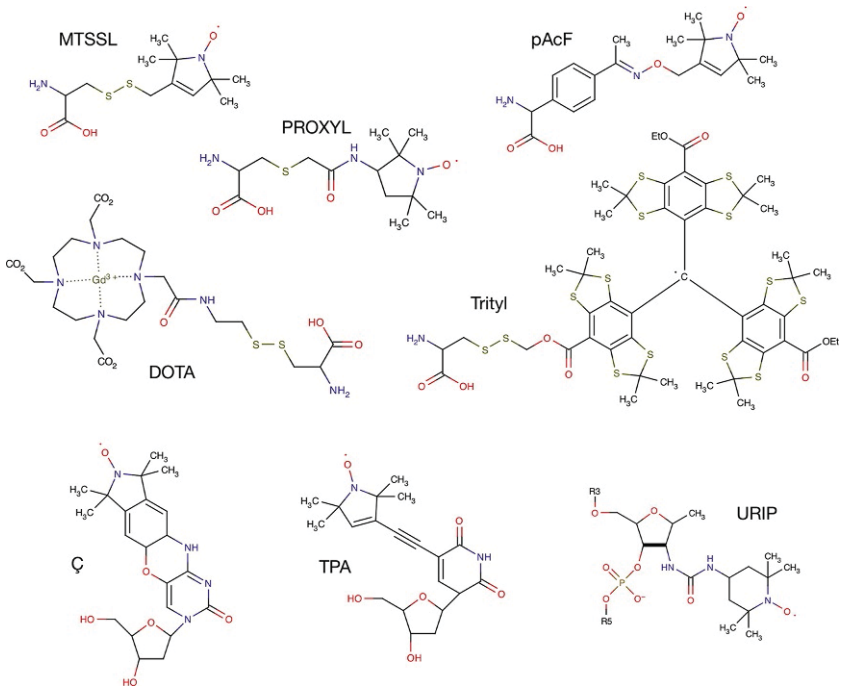


Figure 1 Commonly used spin labels for long-range EPR distance measurements. The MTSSL (Berliner, Grunwald, Hankovszky, & Hideg, 1982), PROXYL (Ramos & Varani, 1998), pAcF (Fleissner et al., 2009), DOTA (Potapov et al., 2010; Song, Meade, Astashkin, & Klein, 2011), and trityl (Reginsson, Kunjir, Sigurdsson, & Schiemann, 2012) labels are used for proteins. The Ç (Cekan, Smith, Barhate, Robinson, & Sigurdsson, 2008), TPA (Piton et al., 2007), and URIP (Edwards, Okonogi, & Robinson, 2001) labels are used for nucleic acids. All labels are shown attached to an amino- or nucleic acid.

dipolar coupling is then converted into distance information, for example, by using the DeerAnalysis software package (Jeschke et al., 2006). The obtained distance distributions can then be used to analyze conformational changes (Hänelt, Wunnicke, Bordignon, Steinhoff, & Slotboom, 2013; Zou et al., 2009) or to reconstruct macromolecular complexes (Kim et al., 2011). Although spin labels are typically small compared to, e.g., FRET labels, (MTSSL has roughly the size of an arginine residue); they still act as a flexible linker between the protein and the spin centre itself (for example, the NO group of MTSSL). This leads to a spatial distribution of the spin center, the size of which strongly depends on the conformational freedom of the spin label at a particular position. Thus, the problem arises, that on the one hand, the PELDOR experiment delivers an accurate distance distribution for the two spin centres, while on the other hand, the exact positions of the underlying distance vectors with respect to the protein are unknown. This problem is commonly solved *in silico*, by modeling conformational ensembles of the spin label onto the protein structure. Different approaches such as the accessible volume approach (Sale, Song, Liu, Perozo, & Fajer, 2005) and the rotamer approach (Polyhach, Bordignon, & Jeschke, 2011) have been developed for this purpose. The former is used by the program mtsslWizard (Hagelueken, Ward, Naismith, & Schiemann, 2012), the latter one by MMM and PRONOX (Hatmal et al., 2011; Polyhach et al., 2011). Briefly, the mtsslWizard, which is the central tool of the mtsslSuite (Hagelueken, Abdullin, Ward, & Schiemann, 2013), attaches a model of the label, e.g., MTSSL (Fig. 1) to the macromolecule and rotates the label around its rotatable bonds. While doing this, such rotamers of the label are rejected, which clash into the macromolecular surface, while allowed conformations are weighted evenly. The result is an ensemble of possible spin label conformations that basically represents the accessible volume of the spin label (Fig. 2). The rotamer approach refines the accessible volume by only allowing such conformations of the spin label, which are in agreement with precalculated rotamer libraries of the free label (Polyhach et al., 2011). In contrast to the accessible volume approach, the rotamers are then also weighted with respect to their internal energy. The performances of the different approaches have been compared in various benchmark studies. Interestingly, the differences are minute, and a prediction error of ~3 Å can be expected (Alexander et al., 2013; Hagelueken et al., 2012; Jeschke, 2013). Once a satisfying model of the spin-labeled macromolecule has been constructed, distances between the modeled conformational ensembles of the spin labels can be calculated and compared to experimental data. In this way, it is, for

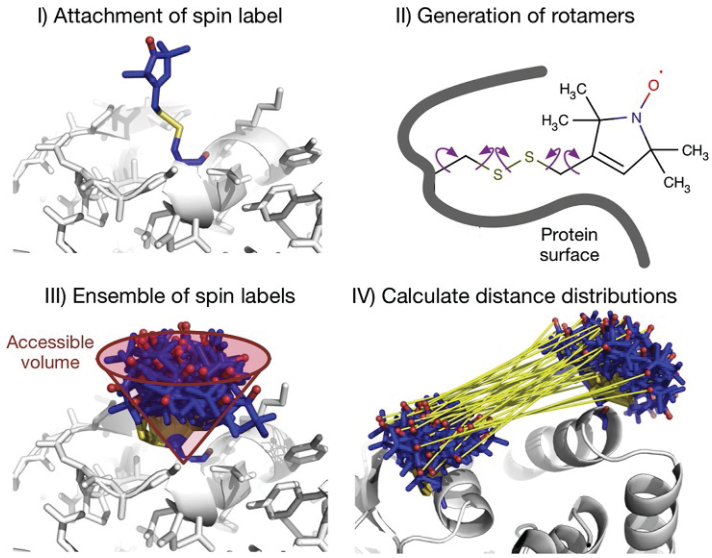


Figure 2 Outline of the *in silico* spin-labeling algorithm used by mtsslWizard. Step (I) The selected label is attached to the macromolecule. Step (II) The program creates rotamers and checks for clashes with the macromolecular surface. Step (III) An ensemble of possible spin label conformations is created. It represents the accessible volume of the spin label. Step (IV) The distance distribution between two spin label ensembles is calculated.

example, possible to decide which molecular model best describes the solution state of the macromolecule.

In this chapter, we describe how to use mtsslWizard, and how the generated models can be utilized for the docking of a macromolecular complex via mtsslDock or for the trilateration of a metal ion by means of mtsslTrilaterate. The mtsslSuite is freely available from the PyMOLWiki (www.pymolwiki.org/) and is tightly integrated into the PyMOL molecular graphics system (www.pymol.org). Many excellent manuals and tutorials for PyMOL are available, and the reader is referenced to the PyMOLWiki (www.pymolwiki.org) as a source for such materials.

2. PLANNING THE PROJECT

The first question during the planning stage of a project is, whether or not PELDOR distance measurements can likely answer the biological question of interest. It should be considered that only such distances that lie within the accessible range for PELDOR spectroscopy can be measured.

This range spans from ~15 to 80 Å (Jeschke, 2012), but longer distances have been measured using fully deuterated proteins (Ward et al., 2010). If the aim of the project is to distinguish between different available models and conformations of a protein, the distance differences have to be large enough to safely distinguish between the models. On top of these requirements, the spin-labeling positions should be located on the molecular surface, preferably in well-structured regions, for example, on α-helices or β-sheets. The reason for the former requirement is that surface sites are easily accessible during the labeling procedure, and the structure of the protein is less likely perturbed when the spin label binds on its surface. The latter requirement is based on the fact that secondary structure elements are usually more rigid than loops, and thus no additional flexibility is added to the inherent flexibility of the spin label.

In this chapter, the heme containing monooxygenase cytochrome P450cam from *Pseudomonas putida* was chosen as an example to introduce the different functions of the mtsslWizard and to show how the program can be helpful to locate optimal labeling sites. P450cam belongs to the large family of cytochrome P450 enzymes (Pylypenko & Schlichting, 2004) and enables the Gram-negative bacterium *P. putida* to metabolize camphor by hydroxylating the compound (Schlichting et al., 2000). The structure and function of P450cam have been thoroughly analyzed in many studies and the conformational changes that are induced by camphor binding have recently been analysed by PELDOR spectroscopy (Stoll et al., 2012).

3. WHERE TO PUT THE SPIN LABEL?

In the following part of this chapter, we demonstrate how mtsslWizard can be used to find optimal labeling positions on a biomolecule (here: P450cam). First, PyMOL is launched, and two coordinate files representing the camphor-bound (PDB-ID: 2CPP; Poulos, Finzel, & Howard, 1987) and apo form (PDB-ID: 3L61; Lee, Wilson, Rupniewski, & Goodin, 2010) of P450cam are loaded, e.g., via "File"->"Load" or by typing "fetch 3L61" and "fetch 2CPP" (Fig. 3). The two structures will appear in the main viewer, and their identifiers will be listed on the right side of the main viewer (Fig. 3, green (light gray in the print version) box). Within PyMOL, the two structures are now referred to as "objects." Any solvent atoms are removed from the objects by typing "remove solvent" into the PyMOL command window (Fig. 3, bottom of blue (dark gray in the print version) box).

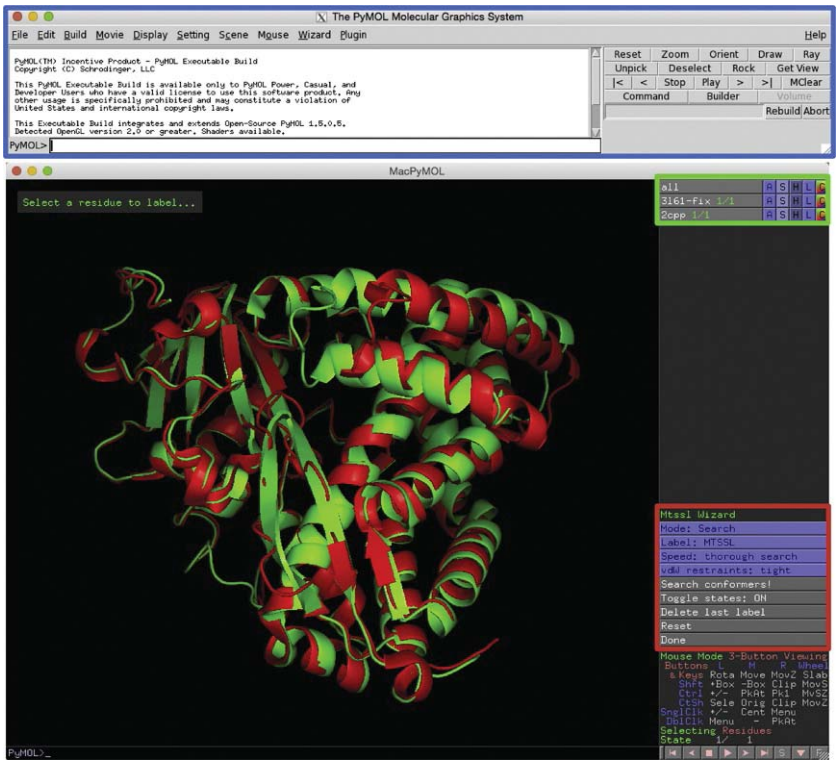


Figure 3 The mtsslWizard GUI. The PyMOL graphical user interface (GUI) with the structures of apo cytochrome (3L61) and camphor-bound cytochrome (2CPP) loaded into the main viewer. The mtsslWizard GUI, the PyMOL command window, and the PyMOL object list are marked by red (gray in the print version), blue (dark gray in the print version), and green (light gray in the print version) boxes, respectively.

This is important, because solvent atoms would otherwise interfere with the *in silico* spin-labeling procedure (see below). For reasons explained below, the two structures are then superimposed by typing "super 2CPP, 3L61." The sequence view of PyMOL ("Display"->"Sequence") reveals that compared to 3L61, some additional amino acid residues (Pos. 90–97: REAGEAYD and Pos. 104: D) have been modeled in the 2CPP structure. These amino acids are disordered and thus missing in the 3L61 structure.

To find conformational changes between the two structures, mtsslWizard can perform an elaborate residue-by-residue structural comparison. For this, both structures must have exactly the same number of amino acid residues, and the residues have to be on the same register. Since the additional residues

in 2CPP are absent in 3L61, mtsslWizard should ignore them for the comparison below. A simple way to achieve this is to copy the additional residues from 2CPP into 3L61 by typing: "create 3L61, 3L61 or 2CPP & resi 90 – 97 + 104." Note that this only works, because the structures were superimposed above. Finally, any alternate side chain conformations have to be removed to avoid interference with the mtsslWizard algorithm. This can be done by typing "remove not (alt "+A)" and "alter all, alt="". The PyMOL object list on the right of the main viewer should now contain the two objects 2CPP and 3L61. All other objects can be either deleted by typing (delete <object name>) or disabled by clicking on the respective label in the object list. The mtsslWizard can then be started via "Wizard"->"mtsslWizard." Its GUI will appear at the bottom right corner of the PyMOL window as shown in Fig. 3 (red (gray in the print version) box).

The program is then switched from the default "Search" mode into the "Distance Map" mode by clicking the "Mode" button. Now, a label (here: MTSSL) is selected form the "Label" menu. Once this settings have been made, the two PyMOL objects that shall be compared are selected by pointing and clicking inside the main viewer. Here, the two objects 3L61 and 2CPP are clicked consecutively. The program will then calculate a Cβ–Cβ-based distance matrix for each of the two "spin-labeled" objects. This means that it iterates over all Cβ atoms of a particular object (for P450cam in the current PyMOL session, each object has 404 residues/Cβ atoms) and calculates the distance to each other Cβ atom in that object. The results are two two-dimensional matrices of 404 × 404 distances, one for 2CPP and one for 3L61. Figure 4 gives a detailed explanation of how such matrices are calculated. To locate conformational changes between the two structures, the two Cβ–Cβ matrices are subtracted, and the absolute values of the differences can be used to construct a difference distance matrix of the two structures (Fig. 4). In principle, the difference distance matrix shows by which distance each Cβ in 2CPP has shifted compared to all Cβ atoms in 3L61 and vice versa (Schneider, 2000). The resulting matrix is color coded between blue (no shift) and red (large shift) (Fig. 5A). In this case, the difference distance matrix reveals a number of interesting facts about the two P450cam structures. First, prominent peaks in the matrix reveal that distance changes of up to 10 Å occur. Second, the P450cam structures may be divided into three rigid bodies: RB1 (Pos. 10–160), RB2 (Pos. 161–250), and RB3 (Pos. 251–410). Magenta boxes in Fig. 5 mark the boundaries of the three rigid bodies. The matrix also shows that

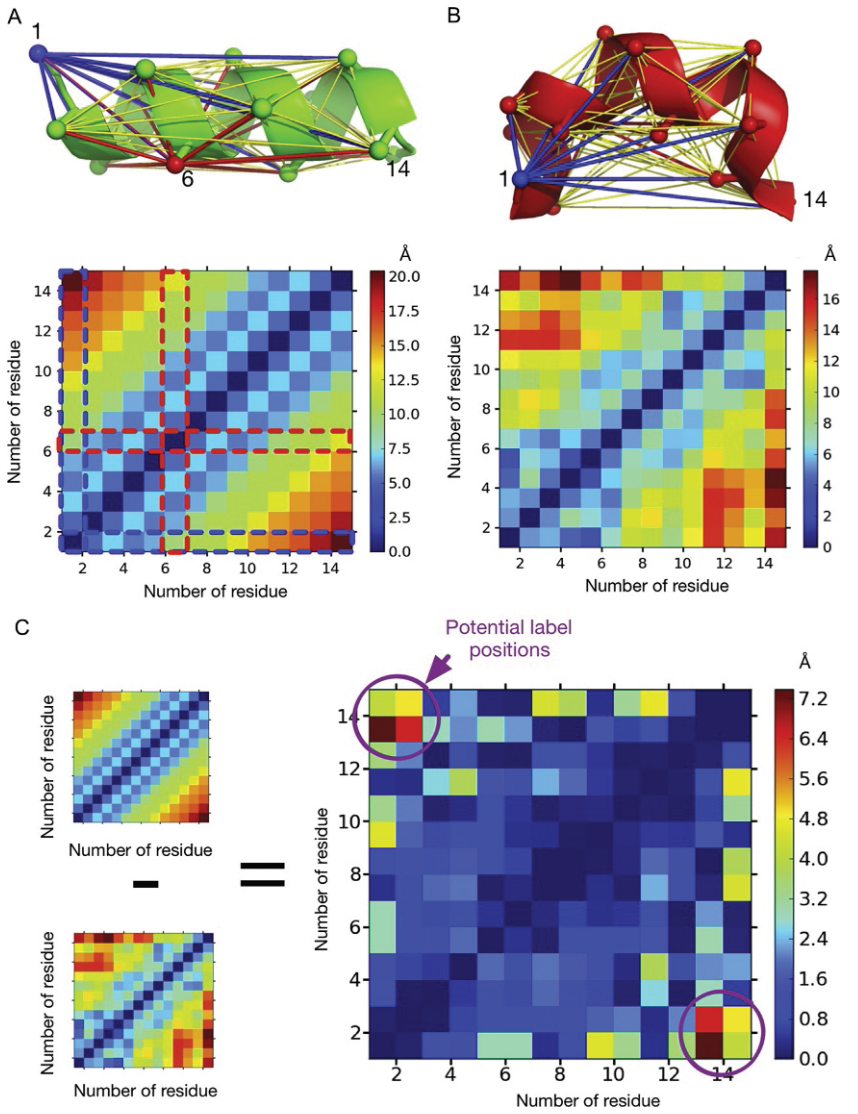


Figure 4 Calculation of distance matrices and difference distance matrices. (A) A protein structure consisting of a single alpha helix with 14 residues is shown as a green cartoon. Spheres indicate the Cβ atoms, and residues 1, 6, and 14 are marked as examples. The network of distance vectors between the Cβ atoms are indicated by yellow lines. The distance vectors for Cβ-1 and Cβ-6 are highlighted in blue and red, respectively. To construct a distance matrix, the absolute values of all distance vectors are calculated and entered into a two-dimensional matrix. The dashed lines indicate were the Cβ-1 and Cβ-6 distances can be found inside the distance matrix. The absolute value

(Continued)

RB2 changes its conformation with respect to both RB1 and RB3. According to the color scale in Fig. 5A, the amplitude of the conformational change ranges up to ~10 Å for particular pairs of residues, and three "conformational hot spots" (HS1, HS2, and HS3) can be defined (Fig. 5A). Pairs of residues within HS1-3 would therefore represent promising candidates for spin label positions.

4. SELECTING OPTIMAL LABELING SITES

The previous step revealed those sites in P450cam, where conformational changes occur when camphor binds (Fig. 5A). The next step is to find labeling positions within these "conformational hot spots" that will yield distances within the PELDOR range. During the calculation started above, mtsslWizard has also estimated the position of *in silico* spin centers on each residue of the 3L61 and 2CPP objects. This operation would take a considerable amount of time (~2 h for each of the two objects) if the default "Search" mode of mtsslWizard was used (Fig. 2). To speed the process up, the "Distance Map" mode uses a faster algorithm that is explained in the following. It is assumed that MTSSL was chosen as a spin label, but the procedure is analog for the other available labels. The algorithm is based on the observation that the "Search" mode of the mtsslWizard distributes the location of the spin center (the part of the spin label which is actually seen by EPR) into an umbrella-like shape. To simulate this, a sphere of randomly distributed trial atoms with a radius of 7.5 Å is constructed and centered on the Cβ atom of the current residue (trial atom sphere) (Fig. 6A). The radius of 7.5 Å corresponds to the distance between the Cβ atom of a "stretched" MTSSL side chain and its spin center, which resides between the N- and O-atoms of the NO group. Now, a second "exclusion sphere" is

Figure 4—Cont'd of each distance is indicated by a color gradient. (B) The structure in panel (A) has undergone a conformational change resulting in the red structure. Because the Cβ-Cβ distances have now changed, the distance matrix of the red structure looks different than the one of the green structure in panel (A). (C) "Conformational hot spots" (see main text) within the two structures can be located by calculating difference distance matrices by subtracting the two matrices shown in panels (A) and (B). The resulting difference distance matrix shows that for this example, the largest conformational changes occur between residues 1 and 13. Thus, these two positions are potential candidates for the attachment of spin labels.

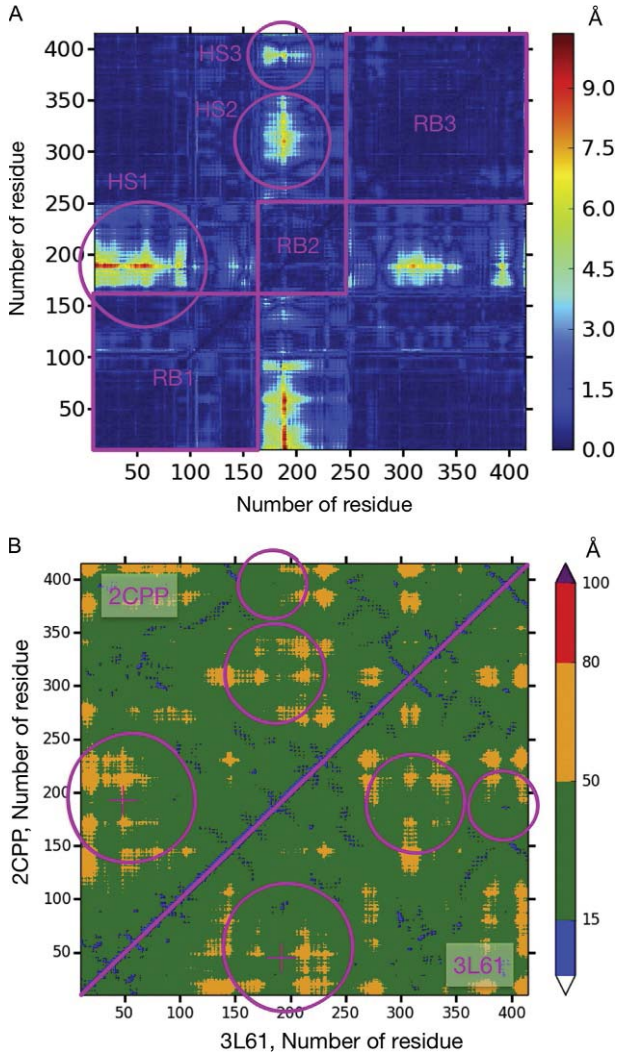


Figure 5 Finding optimal labeling sites on P450cam using distance maps. (A) Difference distance matrix (Cβ–Cβ) of camphor-bound (2CPP) and apo cytochrome P450cam (3L61). The absolute values of the distance changes are represented by a color scale running from blue (no change) to dark red (10 Å change). Rigid bodies 1–3 and conformational hot spots 1–3 (see main text) are marked (RB1-3 or HS1-3). (B) The distance matrices (spin–spin) for 2CPP and 3L61. The matrices are symmetric across their diagonal and were therefore combined into a single matrix for visualization purposes. The upper triangle shows the upper triangle of the 2CPP matrix and the lower triangle the lower triangle of the 3L61 matrix. The diagonal is highlighted in magenta. The locations of HS1-3 (see panel A) are indicated by magenta circles. The distance values are color coded according to their location within the PELDOR distance range. The label positions chosen in the original study (48, 190) are indicated by magenta crosshairs.

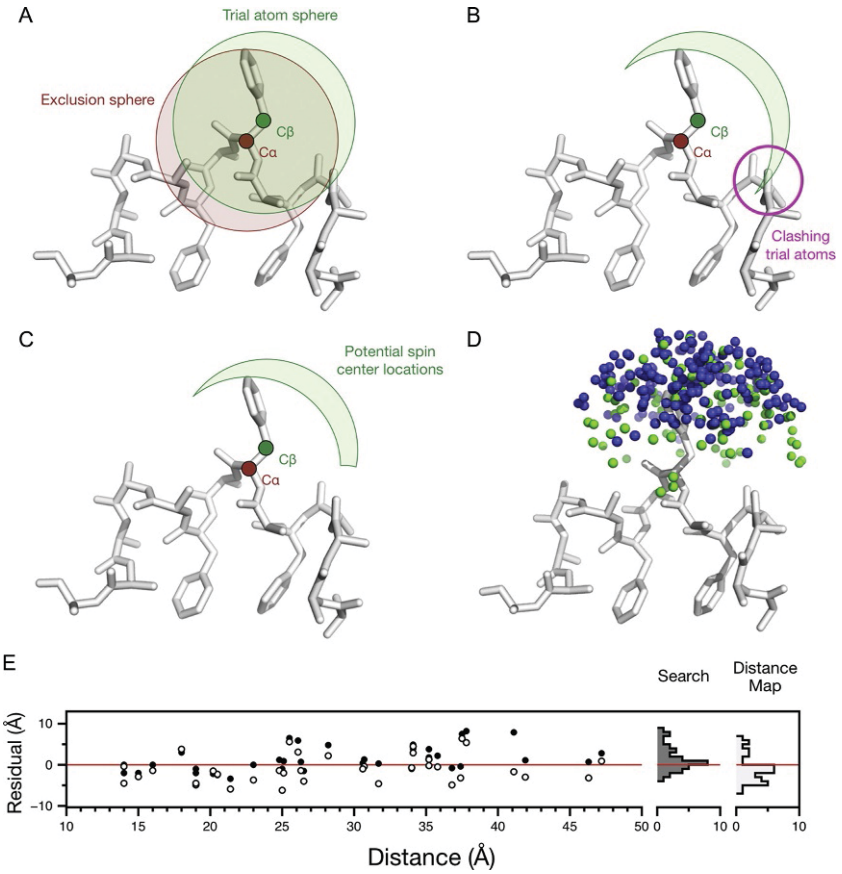


Figure 6 Fast estimation of spin center location by the "Distance Map" mode of mtsslWizard. (A) If MTSSL is selected as the spin label, a sphere of radius 7.5 Å is constructed around the Cβ atom of an amino acid. This "trial sphere" (green (white in the print version)) is filled with randomly distributed trial atoms. The atoms are not shown for clarity. A second exclusion sphere is generated around the Cα atom (red (light gray in the print version)). (B) Any atom in the trial sphere that also resides within the "exclusion sphere" is deleted. Only atoms in the umbrella-like shape (green (white in the print version)) remain. (C) The remaining atoms are checked for clashes with the molecular surface of the protein and clashing trial atoms are removed. The procedure yields a distribution of potential spin center locations. (D) The distribution of trial atoms from panel (C) (green (white in the print version)) is compared to spin center locations that were found using the standard "Search" mode of mtsslWizard (blue (dark gray in the print version)) at the same position. (E) The average value of individual experimental distance distributions between doubly spin-labeled T4-Lysozyme (PDB-ID: 2LZM; Weaver & Matthews, 1987) were compared to the average distances predicted by mtsslWizard. The program was used in the default "Search" mode (black circles). The dark gray histogram on the left illustrates the distribution of the residuals with a mean value of 1.4 Å and a standard deviation of 3.1 Å. The white circles show how the same distances are predicted by the much faster "Distance Map" mode. Here, the mean value is 0.1 Å, and the standard deviation is 3.4 Å.

constructed around the Cα atom. Any atom in the trial atom sphere that also resides within the exclusion sphere is then removed (Fig. 6B). In the next step, any of the trial atoms that are closer than 3.5 Å to the molecular surface are deleted (Fig. 6C, atoms of the currently "labeled" amino acid are ignored). This procedure generates an umbrella-like distribution of potential "spin center" locations around the current amino acid. Figure 6D shows that this distribution is indeed strikingly similar to a distribution produced by the default "Search" mode. The mtsslWizard then calculates the geometric average of the distribution and stores this as a surrogate for the spin center location of this labeling site. For a 400-residue protein and on current computer hardware, this procedure is finished within a few seconds. A benchmark against a previously used test dataset of 36 T4-lysozyme PELDOR distances (Hagelueken et al., 2012) shows that this much faster algorithm performs very similar to the more sophisticated "Search" mode of mtsslWizard (Fig. 6E).

Once this procedure has been completed for all residues in 2CPP and 3L61, distance matrices between the estimated spin center positions are calculated (these matrices are also written out by mtsslWizard). Figure 5B shows a combined spin–spin distance matrix for both the 2CPP and 3L61 objects. Since both matrices are symmetric across their diagonal, they were combined into one matrix for visualization purposes (see Fig. 5). The matrix is color coded to emphasize which of the calculated spin–spin distances lie within the PELDOR window (blue: distance too short (<15 Å), green: optimal distance (15–49 Å), yellow: challenging distance (50–79 Å), red/purple: distance too long (>80 Å)). The locations of the conformational hot spots HS1-3 are indicated by magenta circles in Fig. 5B. The distance matrix reveals that for both structures, several potential label pairs within HS1-3 would generate PELDOR distances in the optimal range between 15 and 49 Å. The authors of the original study chose amino acids S48 and S190 to track the described conformational changes by PELDOR spectroscopy (Stoll et al., 2012). Comparison with Fig. 5A shows that these residues are located in the center of HS1, and the expected distance is within the accessible range for PELDOR spectroscopy (Fig. 5B).

5. WHICH SPIN LABEL SHOULD BE USED?

Provided that suitable labeling sites have been found, the next step is to decide which spin label shall be used. Note that depending on this decision,

the calculations above can be repeated for different labels. In most cases, the nitroxide spin label MTSSL (Fig. 1) is selected because it is experimentally well established, easy to handle, and commercially available (Berliner et al., 1982). Recently, several metal-based spin labels such as the Gd^{3+} labels have become popular (Potapov et al., 2010). These labels have low orientation selectivity and provide high sensitivity. The recently developed trityl- or spirocyclohexyl spin labels have comparably slow relaxation times at room temperature and therefore allow measurements in liquid solution (Kunjir, Reginsson, Schiemann, & Sigurdsson, 2013; Meyer et al., 2015; Reginsson et al., 2012). The mtsslWizard currently contains eight different spin labels for proteins and oligonucleotides, which are shown in Fig. 1. The spin label library of the mtsslWizard is constantly extended, and user requests are always welcome.

Many interesting target molecules have multiple native cysteine residues on their molecular surface. These residues will of course interfere with the labeling reaction and therefore have to be mutated to other amino acids such as serine or alanine. Depending on the number of necessary mutations, this procedure can be time-consuming, and it cannot be predicted, whether the resulting protein will still have the same properties or structure as the native protein. Several workarounds exist for such situations. For example, the noCysHomologs server (http://www-hagelueken.thch.uni-bonn.de/noCysHomologs/form.html) can be used to screen sequence databases for functional homologs, which have fewer or no cysteine residues. Briefly, the server uses the BLAST algorithm (Altschul, Gish, Miller, Myers, & Lipman, 1990) to screen sequence databases for homologs of the protein of interest. These are then simply ranked by the number of contained cysteine residues and compiled into a list. In this way, possible alternative target molecules (if any) can be found within minutes. Another option is the introduction and spin labeling of unnatural amino acids (Fleissner et al., 2009). The mtsslWizard contains the spin-labeled *para*-acetylphenylalanine side chain (Fig. 1) for such cases.

6. PREPARATION OF SPIN-LABELED SAMPLES AND THE PELDOR EXPERIMENT

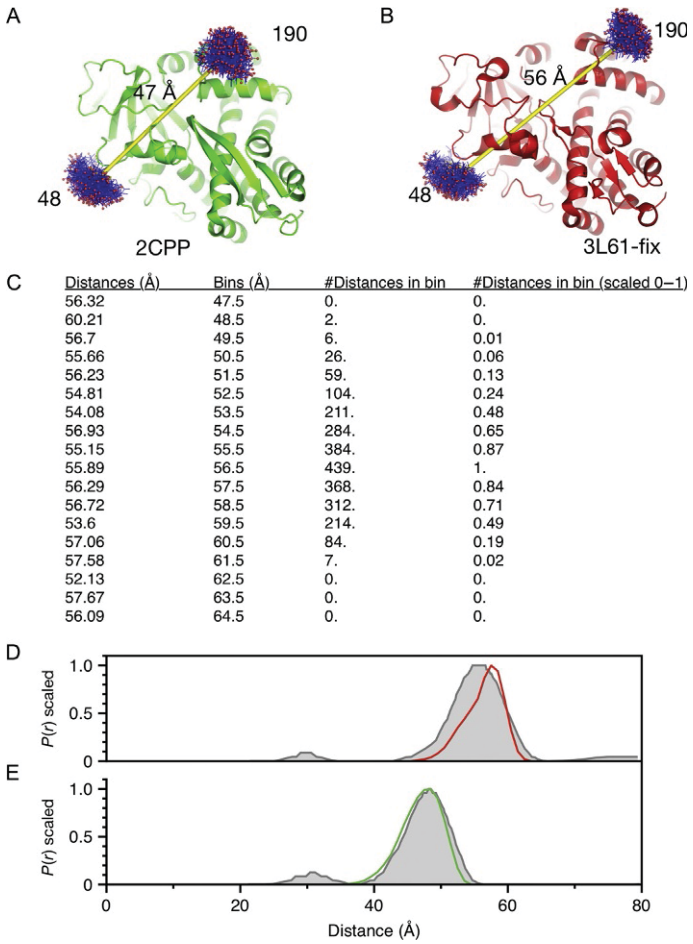
To measure the selected distances, mutants of the protein have to be prepared and spin labeled. After the PELDOR experiment has been performed, the distance data have to be extracted from the PELDOR

time traces. These parts of the procedure are out of the scope of this chapter, and the reader is referred to the many reviews and examples in the literature.

7. TRANSLATING DISTANCE DATA INTO STRUCTURAL INFORMATION

In the following, we assume that positions 48 and 190 of P450cam were spin labeled with MTSSL (see above), and the interspin distances in the presence and absence of camphor were measured using PELDOR. To decide whether the PELDOR data agree with the crystal structures, mtsslWizard will first be used to spin label positions 48 and 190 of both structures *in silico*. For this purpose, the mtsslWizard is switched into its "Search" mode by clicking the "Mode" button (Fig. 3). Note that albeit slower, the "Search" mode will provide more exact label ensembles than the faster "Distance Map" mode. By default, MTSSL is already selected in the "Label" menu. The default setting in the "Speed" menu is "thorough search," which means that the program will evaluate up to 10,000 trial conformations of the spin label until 200 possible conformations have been found. Each trial conformation is checked for clashes with protein atoms and internal clashes. Depending on the setting in the "vdW restraints" menu, the trial conformation is rejected when one (tight restraints) or more than five (loose restraints) clashes are found. Here, a clash is defined as an interatomic distance of $<$3.4 Å (tight restraints) or $<$2.5 Å (loose restraints). To label position 48 in 2CPP, the residue is selected (Tip: it is easiest to select the residue using the sequence view of PyMOL), and the "Search conformers!" button is clicked. The process is then repeated for position 190 and the equivalent residues in 3L61. Each run produces an ensemble of potential spin label conformations attached to the selected site (Fig. 7A and B).

In the next step, the distance distributions between the conformational ensembles are calculated. To do this, mtsslWizard is switched into its "Distance" mode by clicking the "Mode" button (Fig. 3). The program now prompts the user to consecutively click on two spin label ensembles. For the purpose of this example, the ensembles located on 3L61 are clicked. The software then calculates the spin–spin distances between each pair of conformers and writes the distance histogram into a text file. An excerpt from the output file is shown in Fig. 7C. The data can be plotted using any spreadsheet program. The same procedure is repeated for the 2CPP



C

Distances (Å)	Bins (Å)	#Distances in bin	#Distances in bin (scaled 0–1)
56.32	47.5	0.	0.
60.21	48.5	2.	0.
56.7	49.5	6.	0.01
55.66	50.5	26.	0.06
56.23	51.5	59.	0.13
54.81	52.5	104.	0.24
54.08	53.5	211.	0.48
56.93	54.5	284.	0.65
55.15	55.5	384.	0.87
55.89	56.5	439.	1.
56.29	57.5	368.	0.84
56.72	58.5	312.	0.71
53.6	59.5	214.	0.49
57.06	60.5	84.	0.19
57.58	61.5	7.	0.02
52.13	62.5	0.	0.
57.67	63.5	0.	0.
56.09	64.5	0.	0.

Figure 7 *In silico* spin labeling of P450cam. (A) The structure of camphor-bound P450cam is shown as a green (gray in the print version) cartoon. Positions 48 and 190 were spin labeled with mtsslWizard to produce the blue (black in the print version) spin label ensembles. A distance of 47 Å (yellow (light gray in the print version)) was measured between the average coordinate of the spin label ensembles. (B) The structure of apo P450cam is shown analogously to panel (A). Here, a distance of 56 Å was measured between the two spin label ensembles. (C) An example for the results of the "Distance" mode of mtsslWizard. The result file contains four columns and can be imported into any spreadsheet program. Column 1: List of all measured distances between two selected spin label ensembles. Column 2: Histogram bins in 0.5 Å steps running from 0 to 100 Å. Column 3: Number of distances in the corresponding bin of column 2. Column 4: Same as column 3 but scaled between 0 and 1 for comparison purposes. (D) The experimental distance distribution for apo P450cam is shown in gray. The distance distribution calculated by mtsslWizard is shown in red (dark gray in the print version). (E) Same as panel (D) but for P450cam in the presence of camphor.

structure. Figure 7D and E shows plots of the two calculated distance distributions (red (dark gray in the print version) and green (gray in the print version) traces) compared to the experimental data (gray, digitized from Stoll et al., 2012). Thus, in this case, the experimental PELDOR data are in excellent agreement with the two crystal structures. The reader should refer to the original publication (Stoll et al., 2012) for an in-depth analysis. It should also be noted that the experimentally found distance changes of 10 Å are in very good agreement with the difference distance matrix shown in Fig. 5A.

8. CAVEATS OF *IN SILICO* SPIN LABELING

The P450cam example demonstrates that *in silico* spin labeling is very useful for the interpretation of PELDOR data. Nevertheless, some caveats do exist. It should be kept in mind that the mtsslWizard assumes a free rotation of the spin label in its accessible volume and that no allowed conformation is weighted higher than any other allowed conformation. This means that any existing favorable interaction of the spin label with the macromolecular surface can lead to striking differences between experimental and calculated distance distributions (Hagelueken et al., 2012). Also, a rotation of a neighboring side chain (or a misplaced side chain) in one of the model structures can markedly influence the shape of the distance histogram (Florin, Schiemann, & Hagelueken, 2014). Several studies have demonstrated that the average error of mtsslWizard and other *in silico* spin-labeling programs amounts to ~3 Å (Hagelueken et al., 2012; Hirst, Alexander, Mchaourab, & Meiler, 2011; Jeschke, 2013). Interestingly, the error does not seem to depend on the particular *in silico* spin-labeling algorithm used (Hagelueken et al., 2012; Hirst et al., 2011; Jeschke, 2013). Even the very simple "Distance Map" mode of mtsslWizard (explained above) can produce similar results (Fig. 6E). To decrease the error, interactions of the spin label with its molecular surrounding (including the solvent) would have to be explicitly considered by the algorithms (Florin et al., 2014).

As a consequence, it is absolutely necessary to measure multiple PELDOR distances before a particular conformation or structural model is either accepted or rejected. It is also good practice to use other available software packages, such as the MMM package (Polyhach et al., 2011) to cross check whether any drawn conclusion is also supported by a different *in silico* labeling algorithm.

9. DOCKING WITH mtsslDock

Sometimes, two macromolecules are known to form a tight complex, but no structure of the macromolecular assembly is available. In such cases, PELDOR distances can be used as constraints to dock the complex (Hilger, Polyhach, Padan, Jung, & Jeschke, 2007; Kim et al., 2011; Ward et al., 2009; Zoltner et al., 2013). The mtsslSuite contains the mtsslDock program, which uses a genetic algorithm to dock two macromolecules based on a set of experimental PELDOR distances (Hagelueken et al., 2013; Fig. 8).

The program requires models of the spin-labeled macromolecules and the experimental distance data between them as input. The structural models have to be *in silico* spin labeled with mtsslWizard as shown above. Here, we will use the protein complex of rubredoxin and rubredoxin reductase (PDB-ID: 2V3B; Hagelueken et al., 2007) as an example (Fig. 8). The structure is loaded into PyMOL (see above) and two separate PyMOL objects for the reductase (create Protein_A, 2v3b & chain A) and rubredoxin (create Protein_B, 2v3b & chain B) are created. After disabling the original structure (disable 2v3b) and removing solvent atoms (remove solvent), mtsslWizard is used to attach spin labels at positions 65, 178, 90, 249, 306, 97, and 31 of "Protein_A" and 51, 20, 53, 35, 16, 32, and 29 of "Protein_B" (Fig. 8).

MtsslDock is then launched from the PyMOL plugins menu ("Plugins"-> "mtsslDock"), and the "Import" dialog appears as shown in Fig. 9A. The purpose of this dialog is to select which PyMOL objects are docked by the program (here: Protein_A and Protein_B). To do this, Protein_A and Protein_B are selected from the two pull-down menus in the "Select

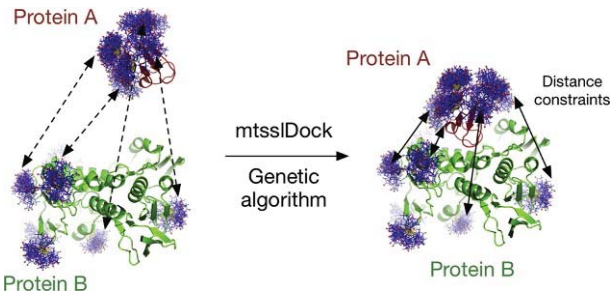
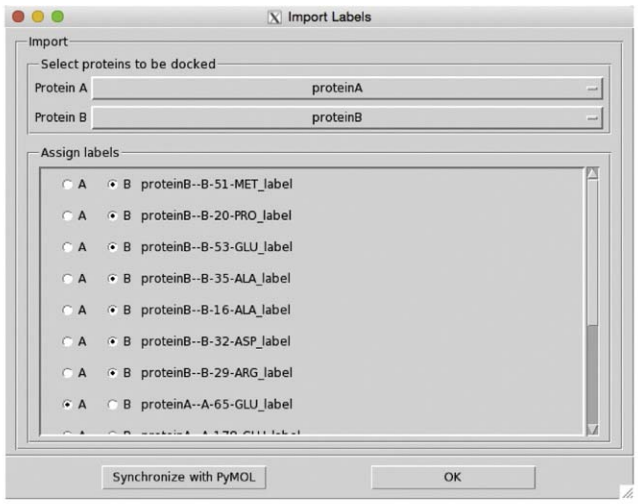
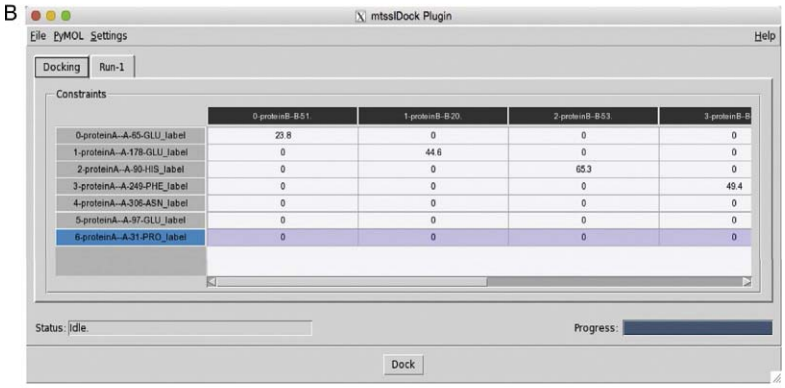


Figure 8 Distance constrained rigid body docking using mtsslDock. The X-ray structure of the protein complex between rubredoxin (red) and rubredoxin reductase (green) is used as an example (Hagelueken et al., 2007). The MTSSL labels (blue) were attached with mtsslWizard. Arrows symbolize available distance constraints.

A



B



C

Pos. on Protein A	Pos. on Protein B	Distance constraint (Å)
65	51	23.8
178	20	44.6
90	53	65.3
249	35	49.4
306	16	45.4
97	32	62.7
31	29	50.6

Figure 9 The MtsslDock GUI. (A) Import dialog of mtsslDock. (B) Docking tab of mtsslDock. (C) Distance constraints between the spin label ensembles shown in panel (A).

proteins to be docked" box, which contains all currently existing objects in PyMOL. At this point, the mtsslDock program does not know whether a particular spin label object in the current PyMOL session belongs to Protein_A or Protein_B; therefore, the label objects have to be manually assigned to the two proteins. The list in the "Assign labels" panel of the

"Import" dialog contains a list of all currently existing spin label objects in PyMOL. To assign a label to either Protein_A or Protein_B, the radio button next to the identifier of the spin label object is clicked. After all selections have been made, the "Import" dialog can be closed by clicking the "OK" button. Using the information from the import dialog, mtsslDock prepares a "Constraints" table inside the "Docking" tab, and the experimental distances and their estimated errors (if available) have to be filled in as shown Fig. 9B. The mean distance and its estimated error have to be separated by a semicolon. If no experimental data exist for a pair of labels, a "0" is entered in the particular cell of the table (Fig. 8B). For this example, the distances shown in Fig. 9C are entered. To start the docking run, the "Dock" button is clicked. The mtsslDock program then creates random trial solutions, which are optimized by its genetic algorithm. The status of the run can be monitored via the "Progress" bar at the bottom right. After the run has finished, a results tab named "Run-1" appears in the GUI (Fig. 10A). Here, all docking solutions are listed in an interactive table on the left. Clicking on one of the solutions centers, the PyMOL session on this object and its docking statistics are listed in a table on the right. In this way, the user can easily check and validate the solution.

Among other things, the statistics table includes the deviation between each experimental constraint and the equivalent distance in the docked complex. The settings that were used for a particular docking run are listed at the bottom left. If necessary, they can be changed via "Settings"->"Docking Settings" and a new run can be started from the "Dock" tab.

With the constraints (Fig. 9C) and label positions (Fig. 9A) from this example, mtsslDock usually finds five solutions of Protein_B, which superimpose well with the crystal structure (Fig. 10B, orange/red (gray/dark gray in the print version)). In addition, five solutions are also found, which match the constraints equally well, but do not form contacts with Protein_A (Fig. 10B, blue (black in the print version)). To find a unique solution, more constraints (and therefore spin labels) would have to be added. While in theory, six intermolecular distances should suffice to find a unique docking solution, the example shows that this is not true in practice. An elaborate benchmark of mtsslDock was performed in the original publication (Hagelueken et al., 2013). It was found that eight or more constraints should be used to find the correct solution with high probability.

The reader should be aware that mtsslDock only performs rigid body docking. Thus, any conformational change upon complex formation is neglected. In practice, this means that solutions where protein B clashes into

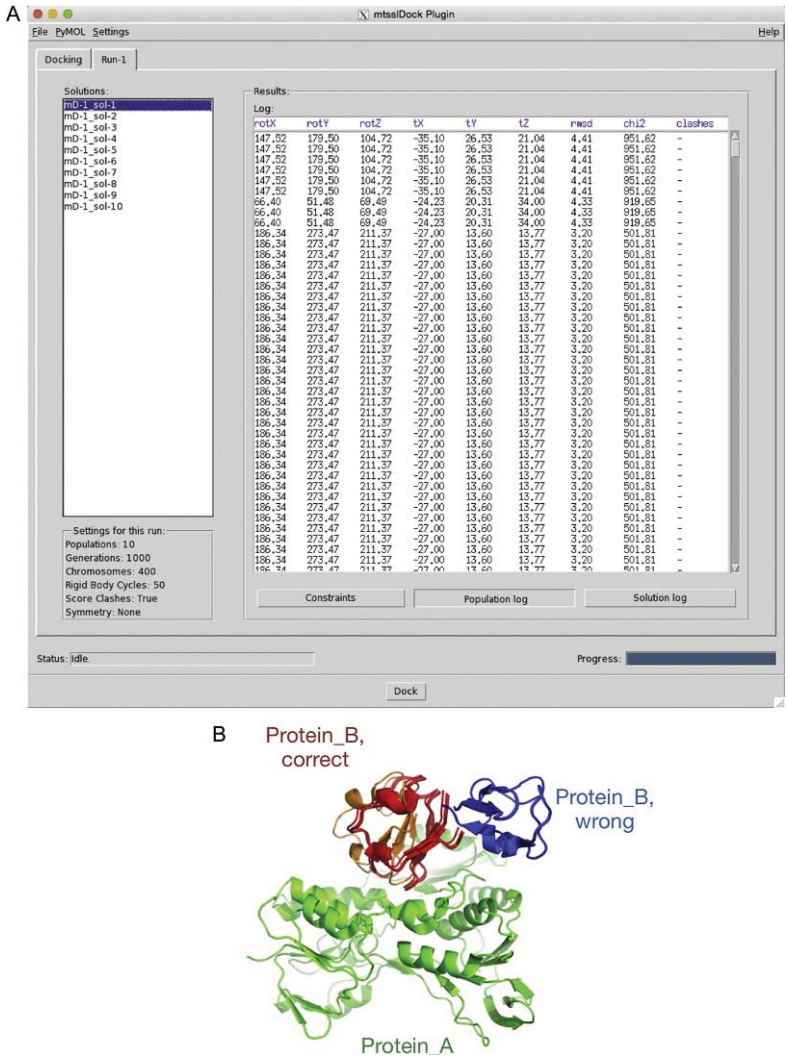


Figure 10 Docking results. (A) The result tab "Run-1." This tab is automatically opened after a docking run has been completed. (B) Typical docking results when the constraints in Fig. 4D are used. The X-ray structure is shown in orange (gray in the print version), the correct docking solutions in red (dark gray in the print version), and the wrong solutions in blue (black in the print version).

protein A, or the distance between the two proteins seems too large, could in principle be caused by conformational changes. Therefore, if only such docking solutions can be found and enough constraints are available, conformational changes should be considered.

10. TRILATERATION WITH mtsslTrilaterate

Akin to the global positioning system, the mtsslTrilaterate program can be used to locate spin centers such as metal ions, cofactors, or spin-labeled substrates in the structure of biomolecules by means of trilateration (Hagelueken et al., 2013). Here, we will use the program to locate the copper ion in the apo structure of azurin (PDB-ID: 1E65; Nar, Messerschmidt, Huber, van de Kamp, & Canters, 1992; Fig. 11A; Abdullin, Florin, Hagelueken, & Schiemann, 2015). The structure of azurin is loaded into PyMOL (see above) and positions 21, 30, 61, 69, 96, and 100 are spin labeled

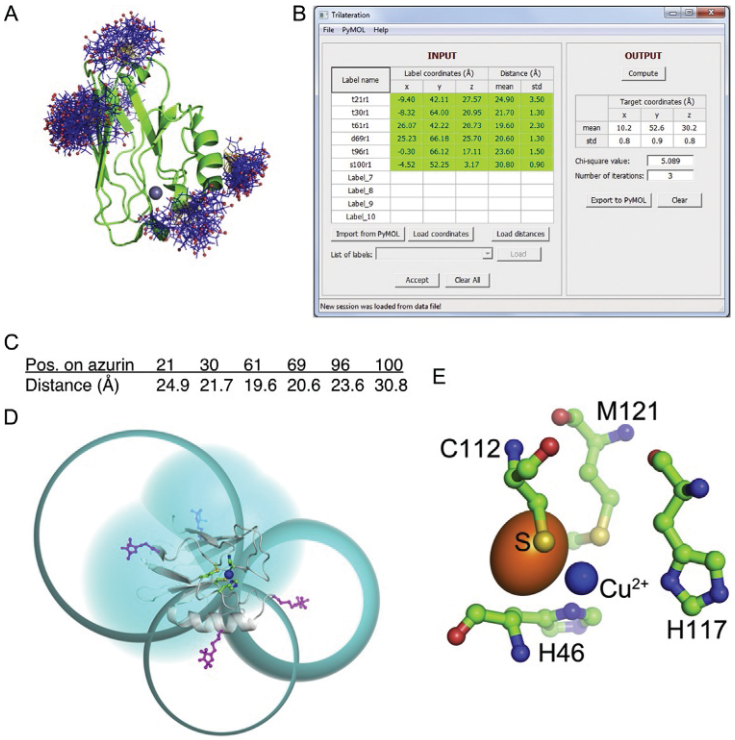


Figure 11 mtsslTrilaterate. (A) The crystal structure of apo azurin (green (light gray in the print version)) with attached MTSSL labels. The approximate position of the Cu^{2+} ion is marked by a gray sphere. (B) mtsslTrilaterate GUI. (C) Cu-nitroxide distances which are used for the example in the main text. (D) Result of the trilateration. The blue (black in the print version) spheres are constructed around the average position of each spin label (pink (gray in the print version)). (E) A close-up of the intersection. The orange (gray in the print version) ellipsoid represents the trilateration result.

with MTSSL via mtsslWizard (see above, Fig. 11A). The mtsslTrilaterate program is then launched via the plugins menu of PyMOL "Plugins"-> "mtsslTrilaterate" and its GUI appears as shown in Fig. 11B. The GUI contains a menu for file operations and a main window. The main window consists of two panels: "Input" and "Output." The "Input" panel represents all input data as a table. Each row of this table is related to one particular spin label. The identifiers of these spin labels are listed in the first column. The following three columns correspond to the average coordinate (x, y, z) of the spin label. These coordinates can be imported directly from PyMOL by selecting the label from the "List of labels" pull-down and clicking the "Import from PyMol button." Here, all labels that were introduced above are now imported into mtsslTrilaterate. The last two columns correspond to the mean value and estimated error of the experimental distances between this particular spin label and the spin center to be localized. The distance data itself can be entered in two different ways: manually or by import of datafiles. The program can, for example, import DeerAnalysis (Jeschke et al., 2006) interspin distance distributions.

For the purpose of this example, the values listed in Fig. 11C are entered. The calculation is then started by clicking the "Compute" button. The program uses a nonlinear least squares optimization to find an optimal solution. A complete discussion of the algorithm is given in the original publication (Hagelueken et al., 2013). Note that at least four distance constraints are needed to achieve a unique trilateration solution. The accuracy of the result will increase if more constraints are used (Abdullin et al., 2015). Once the calculation has finished, the results are displayed in the "Output" panel. The solution of the trilateration problem can be interpreted as the most probable coordinates of the spin center and their standard errors. These data are presented as a table in the "Output" panel (Fig. 11B) and can be visualized graphically inside PyMOL by clicking the "Export to PyMOL" button (Fig. 11D and E). Within PyMOL, the trilateration spheres are displayed around each spin label. The radius of each sphere corresponds to the metal-label distance that was measured for this label. The trilateration solution is located at the intersection of the trilateration spheres and is represented by an orange (gray in the print version) ellipsoid. Its dimensions are proportional to the calculated error of the trilateration solution. In this case, the trilateration result is 2.6 Å off the position which was found by X-ray crystallography but within the binding site (Fig. 11D). The error mostly stems from the *in silico* spin-labeling procedure, but other factors such as the delocalization of the Cu^{2+} spin density contribute as well (Abdullin et al., 2015).

11. CONCLUSION

The mtsslSuite was designed to aid a researcher during the planning and interpretation stages of EPR-long-range distance measurements. Its central tool, the mtsslWizard, can be used to identify optimal spin-labeling positions and to calculate distance distributions between these sites. The *in silico* spin-labeled model structures that are created by mtsslWizard can be used as input models for mtsslTrilaterate and mtsslDock. The latter two programs have been designed to locate spin centers within macromolecules or to assemble macromolecular complexes based on EPR-long-range distance constraints, respectively. When using *in silico* spin-labeling tools, the reader should keep the main caveat of *in silico* spin labeling in mind: its relatively large prediction error of ~3 Å, which can be much larger in extreme cases were the spin label strongly interacts with the protein surface (e.g., Lillington et al., 2011). Until more exact algorithms are available, it is important to stress that any conclusions drawn based on spin label distance distributions should always be supported by measurements between multiple positions. Further, it should be checked whether the same conclusion can be reached if different programs are used (e.g., mtsslWizard Hagelueken et al., 2012, MMM Polyhach et al., 2011, PRONOX Hatmal et al., 2011, or RosettaEPR Hirst et al., 2011).

ACKNOWLEDGMENTS

G.H. and O.S. would like to acknowledge financial support by the SFB 813 of the Deutsche Forschungsgemeinschaft (DFG).

REFERENCES

Abdullin, D., Florin, N., Hagelueken, G., & Schiemann, O. (2015). EPR-based approach for the localization of paramagnetic metal ions in biomolecules. *Angewandte Chemie, International Edition*, *54*, 1827–1831.

Alexander, N. S., Stein, R. A., Koteiche, H. A., Kaufmann, K. W., Mchaourab, H. S., & Meiler, J. (2013). RosettaEPR: Rotamer library for spin label structure and dynamics. *PLoS One*, *8*, e72851.

Altenbach, C., Marti, T., Khorana, H. G., & Hubbell, W. L. (1990). Transmembrane protein structure: Spin labeling of bacteriorhodopsin mutants. *Science*, *248*, 1088–1092.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410.

Bayburt, T. H., Grinkova, Y. V., & Sligar, S. G. (2002). Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters*, *2*, 853–856.

Berliner, L. J., Grunwald, J., Hankovszky, H. O., & Hideg, K. (1982). A novel reversible thiol-specific spin label: Papain active site labeling and inhibition. *Analytical Biochemistry*, *119*, 450–455.

Caffrey, M. (2003). Membrane protein crystallization. *Journal of Structural Biology*, *142*, 108–132.

Cekan, P., Smith, A. L., Barhate, N., Robinson, B. H., & Sigurdsson, S. T. (2008). Rigid spin-labeled nucleoside Ç: A nonperturbing EPR probe of nucleic acid conformation. *Nucleic Acids Research*, *36*, 5946–5954.

Edwards, T. E., Okonogi, T. M., & Robinson, B. H. (2001). Site-specific incorporation of nitroxide spin-labels into internal sites of the TAR RNA; structure-dependent dynamics of RNA by EPR spectroscopy. *Journal of the American Chemical Society*, *123*, 1527–1528.

Endeward, B., Butterwick, J. A., MacKinnon, R., & Prisner, T. F. (2009). Pulsed electron–electron double-resonance determination of spin-label distances and orientations on the tetrameric potassium ion channel KcsA. *Journal of the American Chemical Society*, *131*, 15246–15250.

Fleissner, M. R., Brustad, E. M., Kálai, T., Altenbach, C., Cascio, D., Peters, F. B., et al. (2009). Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 21637–21642.

Florin, N., Schiemann, O., & Hagelueken, G. (2014). High-resolution crystal structure of spin labelled (T21R1) azurin from Pseudomonas aeruginosa: A challenging structural benchmark for in silico spin labelling algorithms. *BMC Structural Biology*, *14*, 16.

Georgieva, E. R., Borbat, P. P., Ginter, C., Freed, J. H., & Boudker, O. (2013). Conformational ensemble of the sodium-coupled aspartate transporter. *Nature Structural & Molecular Biology*, *20*, 215–221.

Gordon-Grossman, M., Zimmermann, H., Wolf, S. G., Shai, Y., & Goldfarb, D. (2012). Investigation of model membrane disruption mechanism by melittin using pulse electron paramagnetic resonance spectroscopy and cryogenic transmission electron microscopy. *The Journal of Physical Chemistry. B*, *116*, 179–188.

Hagelueken, G., Abdullin, D., Ward, R., & Schiemann, O. (2013). mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL. *Molecular Physics*, *111*, 2757–2766.

Hagelueken, G., Ingledew, W. J., Huang, H., Petrovic-Stojanovska, B., Whitfield, C., ElMkami, H., et al. (2009). PELDOR spectroscopy distance fingerprinting of the octameric outer-membrane protein Wza from Escherichia coli. *Angewandte Chemie, International Edition*, *48*, 2904–2906.

Hagelueken, G., Ward, R., Naismith, J. H., & Schiemann, O. (2012). Mtsslwizard: In silico spin-labeling and generation of distance distributions in PyMOL. *Applied Magnetic Resonance*, *42*, 377–391.

Hagelueken, G., Wiehlmann, L., Adams, T. M., Kolmar, H., Heinz, D. W., Tümmler, B., et al. (2007). Crystal structure of the electron transfer complex rubredoxin rubredoxin reductase of Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 12276–12281.

Hänelt, I., Wunnicke, D., Bordignon, E., Steinhoff, H. J., & Slotboom, D.-J. (2013). Conformational heterogeneity of the aspartate transporter GltPh. *Nature Structural & Molecular Biology*, *20*, 210–214.

Hatmal, M. M., Li, Y., Hegde, B. G., Hegde, P. B., Jao, C. C., Langen, R., et al. (2011). Computer modeling of nitroxide spin labels on proteins. *Biopolymers*, *97*, 35–44.

Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K. P., Steinhoff, H. J., et al. (2005). Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of E. coli. *Biophysical Journal*, *89*, 1328–1338.

Hilger, D., Polyhach, Y., Padan, E., Jung, H., & Jeschke, G. (2007). High-resolution structure of a Na+/H+ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements. *Biophysical Journal*, *93*, 3675–3683.

Hirst, S. J., Alexander, N., Mchaourab, H. S., & Meiler, J. (2011). RosettaEPR: An integrated tool for protein structure determination from sparse EPR data. *Journal of Structural Biology*, *173*, 506–514.

Hubbell, W. L., López, C. J., Altenbach, C., & Yang, Z. (2013). Technological advances in site-directed spin labeling of proteins. *Current Opinion in Structural Biology*, *23*, 725–733.

Jensen, M. Ø., & Mouritsen, O. G. (2004). Lipids do influence protein function-the hydrophobic matching hypothesis revisited. *Biochimica et Biophysica Acta*, *1666*, 205–226.

Jeschke, G. (2012). DEER distance measurements on proteins. *Annual Review of Physical Chemistry*, *63*, 419–446.

Jeschke, G. (2013). Conformational dynamics and distribution of nitroxide spin labels. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *72*, 42–60.

Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., et al. (2006). DeerAnalysis2006—A comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance*, *30*, 473–498.

Joseph, B., Korkhov, V. M., Yulikov, M., & Jeschke, G. (2013). Conformational cycle of the vitamin B12 ABC importer in liposomes detected by DEER. *The Journal of Biological Chemistry*, *289*, 3176–3185.

Kim, S., Brandon, S., Zhou, Z., Cobb, C. E., Edwards, S. J., Moth, C. W., et al. (2011). Determination of structural models of the complex between the cytoplasmic domain of erythrocyte band 3 and ankyrin-R repeats 13–24. *The Journal of Biological Chemistry*, *286*, 20746–20757.

Klare, J. P., & Steinhoff, H. J. (2009). Spin labeling EPR. *Photosynthesis Research*, *102*, 377–390.

Krstić, I., Hänsel, R., Romainczyk, O., Engels, J. W., Dötsch, V., & Prisner, T. F. (2011). Long-range distance measurements on nucleic acids in cells by pulsed EPR spectroscopy. *Angewandte Chemie, International Edition*, *50*, 5070–5074.

Kunjir, N. C., Reginsson, G. W., Schiemann, O., & Sigurdsson, S. T. (2013). Measurements of short distances between trityl spin labels with CW EPR, DQC and PELDOR. *Physical Chemistry Chemical Physics*, *15*, 19673–19685.

Lee, Y.-T., Wilson, R. F., Rupniewski, I., & Goodin, D. B. (2010). P450cam visits an open conformation in the absence of substrate. *Biochemistry*, *49*, 3412–3419.

Lillington, J. E. D., Lovett, J. E., Johnson, S., Roversi, P., Timmel, C. R., & Lea, S. M. (2011). Shigella flexneri Spa15 crystal structure verified in solution by double electron electron resonance. *Journal of Molecular Biology*, *405*, 427–435.

Meyer, V., Swanson, M. A., Clouston, L. J., Boratyński, P. J., Stein, R. A., Mchaourab, H. S., et al. (2015). Room-temperature distance measurements of immobilized spin-labeled protein by DEER/PELDOR. *Biophysical Journal*, *108*, 1213–1219.

Milov, A., Salikohov, K., & Shirov, M. (1981). Application of endor in electron-spin echo for paramagnetic center space distribution in solids. *Fizika Tverdogo Tela*, *23*, 975–982.

Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., & Canters, G. W. (1992). Crystal structure of Pseudomonas aeruginosa apo-azurin at 1.85 A resolution. *FEBS Letters*, *306*, 119–124.

Piton, N., Yuguang, M., Stock, G., Prisner, T. F., Schiemann, O., & Engels, J. W. (2007). Base-specific spin-labeling of RNA for structure determination. *Nucleic Acids Research*, *35*, 3128–3143.

Pliotas, C., Ward, R., Branigan, E., Rasmussen, A., Hagelueken, G., Huang, H., et al. (2012). Conformational state of the MscS mechanosensitive channel in solution revealed by pulsed electron–electron double resonance (PELDOR) spectroscopy. *PNAS*, *109*, E2675–E2682.

Polyhach, Y., Bordignon, E., & Jeschke, G. (2011). Rotamer libraries of spin labelled cysteines for protein studies. *Physical Chemistry Chemical Physics*, *13*, 2356–2366.

Potapov, A., Yagi, H., Huber, T., Jergic, S., Dixon, N. E., Otting, G., et al. (2010). Nanometer-scale distance measurements in proteins using Gd3+ spin labeling. *Journal of the American Chemical Society*, *132*, 9040–9048.

Poulos, T. L., Finzel, B. C., & Howard, A. J. (1987). High-resolution crystal structure of cytochrome P450cam. *Journal of Molecular Biology*, *195*, 687–700.

Pylypenko, O., & Schlichting, I. (2004). Structural aspects of ligand binding to and electron transfer in bacterial and fungal P450s. *Annual Review of Biochemistry*, *73*, 991–1018.

Ramos, A., & Varani, G. (1998). A new method to detect long-range protein-RNA contacts: NMR detection of electron-proton relaxation induced by nitroxide spin-labeled RNA. *Journal of the American Chemical Society*, *120*(42), 10992–10993.

Reginsson, G. W., Kunjir, N. C., Sigurdsson, S. T., & Schiemann, O. (2012). Trityl radicals: Spin labels for nanometer-distance measurements. *Chemistry—A European Journal*, *18*, 13580–13584.

Reginsson, G. W., & Schiemann, O. (2011). Pulsed electron–electron double resonance: Beyond nanometre distance measurements on biomacromolecules. *The Biochemical Journal*, *434*, 353–363.

Sale, K., Song, L., Liu, Y.-S., Perozo, E., & Fajer, P. G. (2005). Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *Journal of the American Chemical Society*, *127*, 9334–9335.

Sanders, C. R., & Prosser, R. S. (1998). Bicelles: A model membrane system for all seasons? *Structure*, *6*, 1227–1234.

Schiemann, O., & Prisner, T. F. (2007). Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Quarterly Reviews of Biophysics*, *40*, 1–53.

Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., et al. (2000). The catalytic pathway of cytochrome p450cam at atomic resolution. *Science*, *287*, 1615–1622.

Schmidt, M. J., Borbas, J., Drescher, M., & Summerer, D. (2014). A genetically encoded spin label for electron paramagnetic resonance distance measurements. *Journal of the American Chemical Society*, *136*, 1238–1241.

Schneider, T. R. (2000). Objective comparison of protein structures: Error-scaled difference distance matrices. *Acta Crystallographica. Section D, Biological Crystallography*, *56*, 714–721.

Sessa, G., & Weissmann, G. (1968). Phospholipid spherules (liposomes) as a model for biological membranes. *Journal of Lipid Research*, *9*, 310–318.

Smirnova, I., Kasho, V., Choe, J.-Y., Altenbach, C., Hubbell, W. L., & Kaback, H. R. (2007). Sugar binding induces an outward facing conformation of LacY. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 16504–16509.

Song, Y., Meade, T. J., Astashkin, A. V., & Klein, E. L. (2011). Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers. *Journal of Magnetic Resonance*, *210*, 59–68.

Sonoda, Y., Cameron, A., Newstead, S., Omote, H., Moriyama, Y., Kasahara, M., et al. (2010). Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins. *FEBS Letters*, *584*, 2539–2547.

Stoll, S., Lee, Y.-T., Zhang, M., Wilson, R. F., Britt, D., & Goodin, D. B. (2012). Double electron–electron resonance shows cytochrome P450cam undergoes a conformational change in solution upon binding substrate. *PNAS*, *109*, 12888–12893.

Ward, R., Bowman, A., Sozudogru, E., El-Mkami, H., Owen-Hughes, T., & Norman, D. G. (2010). EPR distance measurements in deuterated proteins. *Journal of Magnetic Resonance*, *207*, 164–167.

Ward, R., Pliotas, C., Branigan, E., Hacker, C., Rasmussen, A., Hagelueken, G., et al. (2014). Probing the structure of the mechanosensitive channel of small conductance in lipid bilayers with pulsed electron–electron double resonance. *Biophysical Journal*, *106*, 834–842.

Ward, R., Zoltner, M., Beer, L., El Mkami, H., Henderson, I. R., Palmer, T., et al. (2009). The orientation of a tandem POTRA domain pair, of the beta-barrel assembly protein BamA, determined by PELDOR spectroscopy. *Structure*, *17*, 1187–1194.

Weaver, L. H., & Matthews, B. W. (1987). Structure of bacteriophage T4 lysozyme refined at 1.7 A resolution. *Journal of Molecular Biology*, *193*, 189–199.

Zoltner, M., Norman, D. G., Fyfe, P. K., El-Mkami, H., Palmer, T., & Hunter, W. N. (2013). The architecture of EssB, an integral membrane component of the type VII secretion system. *Structure*, *21*, 595–603. http://dx.doi.org/10.1016/j.str.2013.02.007.

Zou, P., Bortolus, M., & Mchaourab, H. S. (2009). Conformational cycle of the ABC transporter MsbA in liposomes: Detailed analysis using double electron–electron resonance spectroscopy. *Journal of Molecular Biology*, *393*, 586–597.

[P4] EPR-based approach for the localization of paramagnetic metal ions in biomolecules

Reprinted with permission from

Dinar Abdullin, Nicole Florin, Gregor Hagelueken, Olav Schiemann, "EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules", *Angew. Chem. Int. Ed.* **2015**, *54*, 1827-1831.
DOI: 10.1002/anie.201410396

Own contribution to the manuscript:

- Performing PELDOR measurements and analysis of PELDOR data
- Performing molecular modeling and calculations required for trilateration
- Writing the manuscript

EPR-Based Trilateration

DOI: 10.1002/anie.201410396

EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules**

*Dinar Abdullin, Nicole Florin, Gregor Hagelueken, and Olav Schiemann**

Abstract: *Metal ions play an important role in the catalysis and folding of proteins and oligonucleotides. Their localization within the three-dimensional fold of such biomolecules is therefore an important goal in understanding structure–function relationships. A trilateration approach for the localization of metal ions by means of long-range distance measurements based on electron paramagnetic resonance (EPR) is introduced. The approach is tested on the Cu^{2+} center of azurin, and factors affecting the precision of the method are discussed.*

Metal ions are often crucial for the folding and structural integrity of biomolecules, and they are centers of catalysis in metalloproteins[1] and some ribozymes.[2] To understand how these biomolecules perform their function, it is important to know the location of the metal ions in the biomolecular structure. X-ray crystallography can provide this information with high precision. However, it is not always possible to grow crystals of a biomolecule, and in some cases only the metal-free structure can be crystallized. It is also not always possible to crystallize different conformational states of biomolecules that they adopt during folding or function. Another suitable method is high-resolution nuclear magnetic resonance (NMR) spectroscopy. It can be applied in solution and delivers the structure and dynamics of diamagnetic and paramagnetic biomolecules at an atomistic level, but is limited to biomolecules smaller than about 70 kDa.[3] If the metal ion to be localized is luminescent, also fluorescence resonance energy transfer (FRET) might be applied.[4] Nevertheless, there are still examples where the position of the metal ion in a biomolecule is not determined, for example, copper binding sites in amyloid precursor protein or manganese binding sites in ribozymes.[5] Therefore, complementary to these methods, we report herein a concept for the localization of paramagnetic metal ions by means of electron paramagnetic resonance (EPR) spectroscopy. EPR is very powerful in detecting and characterizing paramagnetic species including paramagnetic metal ions.[6] Compared to the methods above, EPR does not require crystallization of biomolecules, can be applied in solution, is not restricted by the biomolecular size and does not require a reference sample. The idea of the EPR-based approach is similar to the global positioning system (GPS) that can locate an object on the surface of the Earth by measuring distances to a number of GPS satellites: Here, the position of a metal ion in a biomolecular structure can be determined via distance constraints measured between this ion and a number of spin labels attached to the surface of the biomolecule by site-directed spin labelling.[7] The attached spin labels act as reference points in the molecular coordinate system of the biomolecule. Their approximate coordinates can be obtained from computational spin labeling programs, as for example, MMM,[8] mtsslWizard,[9] or PRONOX,[10] and the distance constraints can be measured by EPR techniques, which explore the dipolar interaction between pairs of electron spins.[11] The most common method for distance measurements in the range of 1.5–8 nm is pulsed electron–electron double resonance (PELDOR or DEER).[12]

Recently, a nitroxide-labeled lipid in soybean seed lipoxygenase-1 was localized by trilateration,[13] and two EPR-derived distance constraints were used to narrow down the location of a Cu^{2+} ion in the EcoRI endonuclease–DNA complex.[14] Yet, up until now the precision of the approach has not been tested, and there have been no reports on an EPR-based trilateration of a metal ion in a biomolecule. Metal ions introduce an additional complexity to the approach, because orientation selectivity of the PELDOR experiment and spin density delocalization may have to be taken into account. Therefore, we test in this work the EPR-based trilateration approach on the Cu^{2+} ion in the soluble blue copper protein azurin (Figure 1a). The results are compared with the crystallographic data and the precision of the method and factors influencing it are discussed.

To perform the trilateration, six single cysteine mutants of azurin were expressed, purified, and spin-labeled with (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)methanethiosulfonate (MTSSL): T21C, T30C, T61C, D69C, T96C, and S100C (Figure 1a; Supporting Information). All labeled mutants showed the expected blue color, symmetric gel-filtration peaks at elution volumes corresponding to monomeric azurin (Supporting Information, Figure S1) and continuous-wave X-band EPR spectra confirming the presence of only protein-bound MTSSL spin labels (Supporting Information, Figure S2).

The distance between the Cu^{2+} ion and MTSSL were measured for each of the six mutants by the four-pulse PELDOR technique (Supporting Information). To achieve a large modulation depth, the PELDOR experiments were

[*] D. Abdullin, N. Florin, Dr. G. Hagelueken, Prof. Dr. O. Schiemann
Institute of Physical and Theoretical Chemistry
University of Bonn
Wegelerstrasse 12, 53115 Bonn (Germany)
E-mail: schiemann@pc.uni-bonn.de
Homepage: http://www.schiemann.uni-bonn.de

[**] This work was funded by the German Research Foundation (DFG) with a grant for project A6 within the collaborative research centre SFB813 "Chemistry at Spin Centres".

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201410396.

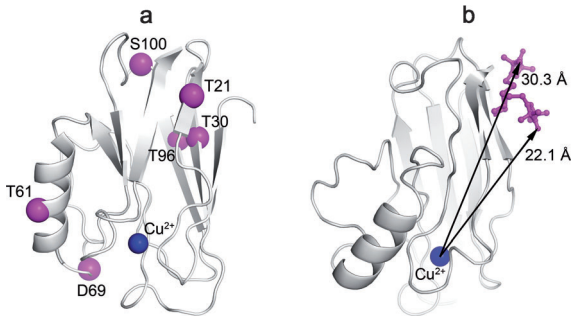


Figure 1. a) The structure of azurin (PDB 1E67) and the chosen mutation labeling sites (pink spheres). b) The crystal structure of the azurin mutant $T21R_1$ (PDB 4BWW) with the two different conformations of MTSSL overlaid.[15] The distance vectors connecting the Cu^{2+} ion with the oxygen atom of MTSSL in $T21R_1$ are shown by arrows.

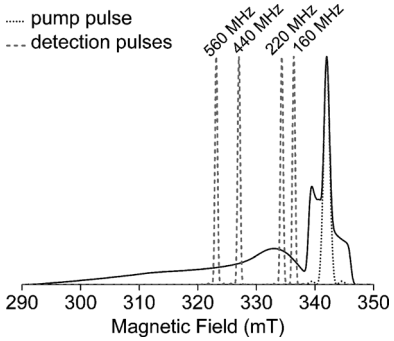


Figure 2. Hahn echo-detected field swept X-band EPR spectrum of azurin mutant $T21R_1$ shown together with the simulated excitation profiles of the 18 ns pump pulse and the 16–32 ns detection pulses.

performed by applying the pump pulse on the maximum of the nitroxide spectrum and the detection pulses at frequencies 160, 220, 440, and 560 MHz higher than the frequency of the pump pulse. Owing to the narrow excitation bandwidth of the pulses (Figure 2), the detection pulses select in dependence of the frequency-offset different spectral components of the copper spectrum and therefore different orientations of the dipolar distance vector connecting the copper ion with the nitroxide.[16] This orientation selection has to be taken into account in the analysis.

The background-subtracted PELDOR time traces of the six azurin mutants are compiled in Figure 3a. All of the time traces exhibit prominent dipolar oscillations and have modulation depths similar to the modulation depths observed for Cu^{2+}/nitroxide model systems.[16] The period of the oscillations and the modulation depth vary between the different frequency offsets, owing to the orientation selectivity and the limited bandwidth of the EPR resonator, which leads to longer pump pulses and therefore smaller modulation depths at larger offsets.

Owing to the orientation selectivity, distance distributions cannot be accurately extracted from a single PELDOR time trace. Therefore, we applied three alternative methods of data analysis. In the simplest method, the distances are estimated from the perpendicular component ($\nu_\perp$) of the Pake doublet, which is obtained after Fourier transformation (FFT) of the

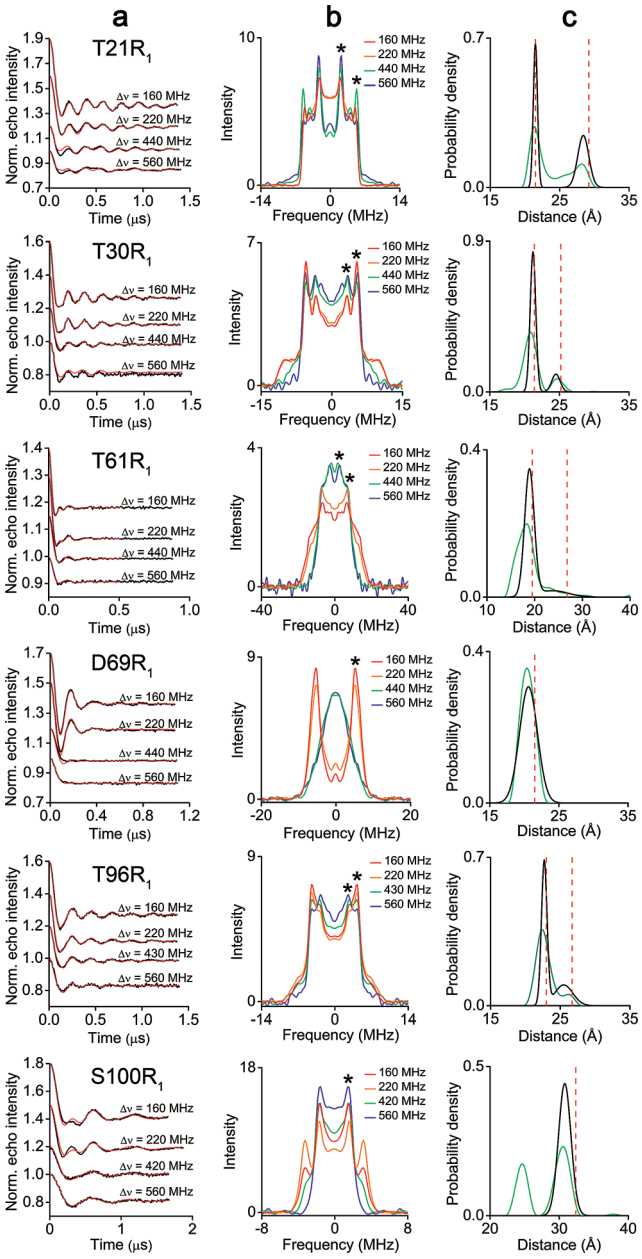


Figure 3. PELDOR data of the six azurin mutants. a) The background corrected PELDOR time traces (black) overlaid with their fits obtained from the PeldorFit program (red). b) The FFTs of the time traces. The positions of the $\nu_\perp$ frequencies are marked by asterisks. c) The distance distributions obtained with FFT (red dashed), DeerAnalysis (green), and PeldorFit (black).

PELDOR time traces and averaging over the four FFTs for each mutant (Figure 3b). In the second method, the time traces recorded with different frequency offsets are summed, again to reduce the orientation selectivity, and then analyzed by Tikhonov regularization as implemented in the DeerAnalysis program.[17] The accuracy of the distance distributions obtained by this approach depends on the efficiency of the orientation averaging. This efficiency is usually small for broad EPR spectra of metal ions. The third method uses the PeldorFit program,[18] which fits the four PELDOR time

Table 1: PELDOR-derived Cu^{2+}-MTSSL distances and their predictions from the molecular simulations.

Mutant	FFT [Å][a]	FFT, average [Å][a]	DeerAnalysis [Å][b]	DeerAnalysis, average [Å][b]	PeldorFit [Å][b]	PeldorFit, average [Å][b]	mtsslWizard [Å][c]
$T21R_1$[d]	21.5, 5.1; 29.2, 1.3	25.3, 3.7	21.7, 1.0; 27.3, 1.4	23.8, 2.9	21.5, 0.3; 28.4, 0.8	24.9, 3.5	24
$T30R_1$	21.4, 2.4; 25.2, 1.3	23.3, 1.9	20.6, 1.1; 24.8, 0.8	21.3, 1.9	21.2, 0.4; 24.5, 0.6	21.7, 1.3	25
$T61R_1$	19.5, 1.9; 26.8, 6.1	23.1, 4.5	18.0, 1.7; 25.3, 2.6	18.9, 3.2	18.9, 1.0; 23.8, 3.2	19.6, 2.3	21
$D69R_1$	21.5, 1.3	21.5, 1.3	20.5, 1.1	20.5, 1.1	20.6, 1.3	20.6, 1.3	20
$T96R_1$	23.0, 1.7; 26.7, 3.0	24.8, 2.4	22.6, 1.0; 26.1, 0.7	23.0, 1.5	22.7, 0.4; 25.5, 1.3	23.6, 1.5	24
$S100R_1$	32.4, 7.8	32.4, 7.8	30.4, 1.0	30.4, 1.0	30.8, 0.9	30.8, 0.9	33

[a] Distances are given in the form of mean value, error (for calculation of errors, see the Supporting Information). [b] Distances are given in the form of mean value, standard deviation. [c] Distances averaged over all generated MTSSL conformers are given. [d] Distances calculated from the crystal structure are 22.1 Å and 30.3 Å.

traces with one set of parameters and provides a geometric model of the spin pair, taking orientation selection into account (Supporting Information).

The distances obtained by these three methods are shown in Figure 3c and Table 1. Interestingly, bimodal distributions were found with all three methods for mutants $T21R_1$, $T30R_1$, $T61R_1$, and $T96R_1$. Since the wild-type azurin cannot be labeled to any significance (Supporting Information, Figure S2), a second labeling site as reason for the second distance can be excluded. At the same time, the pronounced bimodal distance distribution found for mutant $T21R_1$ fits to the crystal structure of this mutant, revealing the presence of two MTSSL conformations (Figure 1b).[15] Moreover, the Cu^{2+}–MTSSL distances calculated for these conformations are in very good agreement with the distances determined by PELDOR (Table 1). In the three other mutants, the second distance peak appears only as a shoulder, but also for these cases the presence of two MTSSL conformations is the most likely explanation. Note that the presence of distinct MTSSL conformations has been found also for other proteins.[19]

To convert the obtained distances into single-value distance constraints required for the trilateration, the corresponding mean distances and standard deviations were calculated (Table 1). Along with the distance constraints, the trilateration requires the coordinates of MTSSL within the molecular frame of azurin. As the reference coordinate system, the crystal structure of the metal-free (*apo*) form of azurin (PDB 1E65) is used. Models of MTSSL attached to the protein surface are created within this coordinate system by means of the program mtsslWizard (Supporting Information, Figure S8). Owing to the intrinsic flexibility of MTSSL, many conformers are created for each site. Therefore, to perform the trilateration, the coordinates of MTSSL at each individual site were averaged over all conformers.

Finally, the trilateration problem is solved by means of the program mtsslTrilaterate[20] using the distance constraints from PeldorFit and the MTSSL coordinates from mtsslWizard (Supporting Information). The position of the Cu^{2+} ion obtained from the calculations is shown within the azurin structure in Figure 4a. It is drawn as an ellipsoid with a center set to the most probable coordinates and semi-principal axes set to the doubled standard errors (2σ, 95% confidence level). To compare this solution with the crystallographic position of the Cu^{2+} ion, the structure of metal-bound azurin was superimposed onto the *apo* structure. The superposition

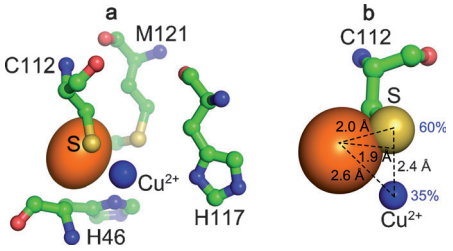


Figure 4. The trilateration of the Cu^{2+} ion in azurin overlaid with a) the crystallographic position of Cu^{2+} and b) its spin density distribution (the spin densities on the Cu and S atoms are shown in percent). The calculated 2σ area of the Cu^{2+} ion location is depicted by an orange ellipsoid, and the corresponding crystallographic position is shown by a blue sphere. The amino acid site chains comprising the binding site are shown as ball-and-stick models.

revealed high similarities between both structures (RMSD of 0.22 Å for 741 atoms), meaning that the binding site is preserved in both structures. As one can see from Figure 4a, the trilateration result places the Cu^{2+} ion within the binding site comprised of residues H46, H117, C112, and M121. Yet, a small shift of the calculated Cu^{2+} position from the corresponding crystallographic site towards the residue C112 is obtained. The value of this shift amounts to 2.6 Å and is the error of the trilateration result. However, it has to be kept in mind that PELDOR is sensitive to the spin density distribution and about 60% of the spin density is located on the sulfur atom of residue C112 and only 35% reside on the Cu^{2+} ion.[21] As a consequence, the obtained distances correspond to the distribution of the spin density, the midpoint of which is shifted from the Cu^{2+} ion towards the sulfur of C112 by 1.5 Å (Figure 4b). This explains the shift of the obtained solution towards residue C112. The offset between the center of the trilateration result and the weighted average position of the spin density is 1.9 Å. In principle, the spin-density distribution can be taken into account already at the stage of the distance calculation,[16b,c] but if the structure of the binding site is not known, this is not possible to achieve.

The remaining error arises from errors in the distance constraints and the MTSSL coordinates. As the exact values of the distances and coordinates are unknown, these errors cannot be easily quantified and separated from each other. Nevertheless, to estimate the error of the MTSSL coordinates, the experimental distances obtained from the PeldorFit

analysis were compared to the corresponding values predicted by mtsslWizard for the crystal structure of metal-bound azurin (Table 1). This comparison revealed an average deviation of 1.5 Å, which is in the same range as found in benchmark studies of the computational spin-labelling programs.[8,9] The error of the PeldorFit analysis is below 1 Å for the distances corresponding to the prominently populated MTSSL conformer, but can increase up to 3 Å for the weakly populated conformer (Supporting Information, Table S3). Furthermore, to estimate the error that the different PELDOR data analysis methods introduce into the distance constraints, the trilateration was performed alternately with distances obtained from the three different methods outlined above (Table 1). Using DeerAnalysis yields a solution deviating by 2.7 Å from the Cu^{2+} crystallographic site and thus having a similar precision as the solution found for the PeldorFit-derived distances. With the distances derived from the FFT approach, this difference increases up to 4.2 Å (Supporting Information, Figure S9). This result shows the importance of an accurate PELDOR data analysis.

Another parameter that influences the precision of trilateration is the number of spin label sites/distance constraints. To reveal this dependence, the trilateration was performed for 6, 5, and 4 constraints (Supporting Information, Figure S10). The results of the experiment show a gradual decrease of the error for the calculated Cu^{2+} position from 4.4 Å for 4 constraints down to 2.6 Å for 6 constraints. Moreover, the average 2σ of the calculated Cu^{2+} coordinates decreases from 2.9 Å down to 1.7 Å, and thus rendering the solution more localized. As a result, increasing the number of constraints improves the precision of trilateration, but requires longer experimental time. Beyond the number of used spin labels, their position with respect to the metal binding site and with respect to each other also affects the result of trilateration. The most reliable solution is achieved when all spin labels are evenly distributed on the protein surface.

Furthermore, the quality of the protein structure or its model has an impact on the trilateration result. To test this, we created a homology model of azurin using the structure of the related blue copper protein auracyanin as starting structure (PDB 2AAN, 30% identical amino acids, RMSD of 1.26 Å between both structures) for the SWISS-MODEL server[22] (Supporting Information). This test revealed that the calculated position of the Cu^{2+} ion is again very close to the binding site (error of 2.0 Å, $2\sigma = 1.6$ Å; Supporting Information, Figure S11) and shows that minor imperfections in the model structure are tolerated.

In conclusion, the EPR-based trilateration approach for metal ions in biomolecules has been successfully tested for the Cu^{2+} center in azurin. The calculated position of the Cu^{2+} ion is within the metal binding site and has a difference to the corresponding crystallographic position of 2.6 Å. The obtained precision of the method is affected by the spin density distribution of the metal center, the accuracy of the MTSSL models, and the number and position of the distance constrains. The error related to the spin density distribution might be further reduced by combining the trilateration approach with electron nuclear double-resonance experiments. This approach can be transferred to other metal ions, including high-spin metal ions such as Fe^{3+} or Mn^{2}. It can be also useful for correlating the position of metal ions found by X-ray crystallography with EPR spectroscopically identified metal ions.[5b] And last but not least, it can be applied for localizing metal ions in folding intermediates or in conformational states of biomolecules. In these cases, models of the biomolecular structure, perhaps constraint by nitroxide–nitroxide distance measurements, would be a prerequisite.

Received: October 23, 2014
Published online: December 17, 2014

Keywords: binding sites · copper · EPR spectroscopy · spin labels · triangulation

[1] *Handbook of Metalloproteins*, *Vol. 4–5* (Ed.: J. Finkelstein), Wiley, Chichester, **2011**.
[2] *Ribozymes and RNA Catalysis* (Eds.: D. M. J. Lilley, F. Eckstein), RSC, Cambridge, **2008**.
[3] a) A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *Curr. Opin. Chem. Biol.* **2005**, *9*, 501–508; b) I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *Dalton Trans.* **2008**, 3782–3790.
[4] a) A. Muschielok, J. Andrecka, A. Jawhari, F. Bruckner, P. Cramer, J. Michaelis, *Nat. Methods* **2008**, *5*, 965–971; b) J. W. Taraska, M. C. Puljung, W. N. Zagotta, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16227–16232.
[5] a) K. J. Barnham, W. J. McKinstry, G. Multhaup, D. Galatis, C. J. Morton, C. C. Curtain, N. A. Williamson, A. R. White, M. G. Hinds, R. S. Norton, K. Beyreuther, C. L. Masters, M. W. Parker, R. Cappai, *J. Biol. Chem.* **2003**, *278*, 17401–17407; b) N. Kisseleva, S. Kraut, A. Jäschke, O. Schiemann, *HFSP J.* **2007**, *1*, 127–136.
[6] a) *Paramagnetic Resonance of Metallbiomolecules* (Ed.: J. Telser), American Chemical Society, Washington, DC, **2003**; b) S. van Doorslaer in *Electron Paramagnetic Resonance*, *Vol. 21* (Ed.: B. C. Gilbert), RSC, London, **2008**, pp. 162–183; c) D. Goldfarb in *Structure and Bonding*, *Vol. 152* (Eds.: C. R. Timmel, J. R. Harmer), Springer, Berlin/Heidelberg, **2014**, pp. 163–204.
[7] a) C. Altenbach, T. Marti, H. Khorana, W. Hubbell, *Science* **1990**, *248*, 1088–1092.
[8] Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356–2366.
[9] G. Hagelueken, R. Ward, J. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377–391.
[10] M. M. Hatmal, Y. Li, B. G. Hegde, P. B. Hegde, C. C. Jao, R. Langen, I. S. Haworth, *Biopolymers* **2012**, *97*, 35–44.
[11] a) O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1–53; b) G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446; c) P. P. Borbat, J. H. Freed in *Structure and Bonding*, *Vol. 152* (Eds.: C. R. Timmel, J. R. Harmer), Springer, Berlin/Heidelberg, **2014**, pp. 1–82.
[12] a) A. D. Milov, K. M. Salikhov, M. D. Shirov, *Fiz. Tverd. Tela* **1981**, *23*, 975–982; b) R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H. W. Spiess, *Angew. Chem. Int. Ed.* **1998**, *37*, 2833–2837; *Angew. Chem.* **1998**, *110*, 2993–2998.
[13] B. J. Gaffney, M. D. Bradshaw, S. D. Frausto, F. Wu, J. H. Freed, P. Borbat, *Biophys. J.* **2012**, *103*, 2134–2144.
[14] Z. Yang, M. R. Kurpiewski, M. Ji, J. E. Townsend, P. Mehta, L. Jen-Jacobson, S. Saxena, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E993–E1000.
[15] N. Florin, O. Schiemann, G. Hagelueken, *BMC Struct. Biol.* **2014**, *14*, 16.

[16] a) E. Narr, A. Godt, G. Jeschke, *Angew. Chem. Int. Ed.* **2002**, *41*, 3907 – 3910; *Angew. Chem.* **2002**, *114*, 4063 – 4066; b) B. E. Bode, J. Plackmeyer, T. F. Prisner, O. Schiemann, *J. Phys. Chem. A* **2008**, *112*, 5064 – 5073; c) B. E. Bode, J. Plackmeyer, M. Bolte, T. F. Prisner, O. Schiemann, *J. Organomet. Chem.* **2009**, *694*, 1172 – 1179; d) J. Sarver, K. I. Silva, S. Saxena, *Appl. Magn. Reson.* **2013**, *44*, 583 – 594.

[17] a) G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473 – 498; b) A. Godt, M. Schulte, H. Zimmermann, G. Jeschke, *Angew. Chem. Int. Ed.* **2006**, *45*, 7560 – 7564; *Angew. Chem.* **2006**, *118*, 7722 – 7726.

[18] D. Abdullin, G. Hagelueken, R. I. Hunter, G. M. Smith, O. Schiemann, *Mol. Phys.* DOI: 10.1080/00268976.2014.960494.

[19] a) G. Hagelueken, W. J. Ingledew, H. Huang, B. Petrovic-Stojanovska, C. Whitfield, H. ElMkami, O. Schiemann, J. H. Naismith, *Angew. Chem. Int. Ed.* **2009**, *48*, 2904 – 2906; *Angew. Chem.* **2009**, *121*, 2948 – 2950; b) C. J. López, Z. Yang, C. Altenbach, W. L. Hubbell, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4306 – E4315.

[20] G. Hagelueken, D. Abdullin, R. Ward, O. Schiemann, *Mol. Phys.* **2013**, *111*, 2757 – 2766.

[21] a) M. van Gastel, J. W. A. Coremans, H. Sommerdijk, M. C. van Hemert, E. J. J. Groenen, *J. Am. Chem. Soc.* **2002**, *124*, 2035 – 2041; b) C. Remenyi, R. Reviakine, M. Kaupp, *J. Phys. Chem. B* **2007**, *111*, 8290 – 8304.

[22] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics* **2006**, *22*, 195 – 201.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules**

*Dinar Abdullin, Nicole Florin, Gregor Hagelueken, and Olav Schiemann**

anie_201410396_sm_miscellaneous_information.pdf

Supporting Information

Table of contents

1. Cloning, protein expression, purification and spin labeling of azurin mutants

The gene for azurin (*azu*, PA4922) was PCR amplified from genomic *Pseudomonas aeruginosa* DNA using the PCR primers 5'-TTATAACCATGGCCGAGTGCTCGGTGG-3' and 5'-TTATAAGAGCT CTCACTTCAGGGTCAGGGTG-3'. The resulting PCR product did not contain the coding region for the N-terminal signal peptide of azurin (residues -20 - 0), so that the target protein (residues 1 - 128 of azurin) could be expressed intracellularly in *E. coli*. The PCR product was then cloned into the vector pEHISGSTTEV (Huanting Liu, Biomedical Sciences Research Center, University of St Andrews, UK) via restriction enzymes NcoI and SacI, resulting in an expression construct with a TEV cleavable N-terminal His_6-GST (glutathione S-transferase) tag. Cysteine single mutants T21C, T30C, T61C, D69C, T96C and S100C were introduced into this construct using PCR.[1] The resulting constructs were each transformed into *E. coli* Rosetta cells. A single colony was picked and grown over night in 50 ml of 2xYeast-Trypton (2YT) media supplemented with 50 µg/ml kanamycin and 17 µg/ml chloramphenicol with shaking at 37°C. On the following day, 1 l of 2YT medium with 50 µg/ml kanamycin and 17 µg/ml chloramphenicol were inoculated with 20 ml of the overnight culture and grown to an OD_{600} of 1.0. Protein expression was then induced by addition of 0.3 mM iso-propoyl-beta-thiogalactoside (IPTG). The protein expression was allowed to proceed for 3 h at 37 °C with shaking at 200 rpm. The cells were then harvested by centrifugation at 2800 g, resuspended in 50 ml of lysis buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 30 mM Imidazol) and lysed with a cell disrupter at 30 kPsi (Constant Systems). Cell debris and insoluble proteins were spun down at 32.000 g for 15 min at 4°C. The soluble fraction was mixed with 1.5 ml Ni-NTA resin (Pierce, pre-equilibrated in lysis buffer) and incubated for 1 h at 4 °C with shaking. The resin was washed with 100 ml of lysis buffer, followed by 50 ml of lysis buffer supplemented with 1 mM tris(2-carboxyethyl)phosphine (TCEP) to reduce the introduced cysteine residues. The reducing agent was then quickly removed by washing the column with 50 ml of lysis buffer, directly followed by addition of 15 ml elution buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 M Imidazol) containing 0.7 mM of MTSSL. A large excess (~20x) of MTSSL was used, since the GST-tag of the expression construct also contained four cysteine residues. The labeling reaction was transferred to dialysis tubing and dialyzed over night against 5 l of dialysis buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl). On the next day, 4 mg TEV protease were added to the sample to cleave the GST-tag. The cleavage reaction was incubated for 3 h at room temperature. The sample was then concentrated to a volume of 2 ml, supplemented with 1 mM $CuCl_2$ and loaded onto a Superdex200 16/60 column (GE)

equilibrated with gel filtration buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl). Labeled monomeric azurin eluted at a volume of ~100 ml and had an intense blue color. For the H_2O/D_2O exchange, the protein was concentrated to 100 µl and filled up to 2 ml with the deuterated buffer (0.1 M TES pH 7.5, 100 mM NaCl in D_2O). This exchange procedure was repeated three times. For cryoprotection, 50% v/v deuterated ethyleneglycol was added, so that the final protein concentration was ~250 µM. The sample was transferred to a standard EPR tube and frozen in liquid N_2 until further use.

2. EPR measurements

All EPR measurements were carried out using a Bruker ELEXSYS E580 EPR spectrometer. Measurements at low temperatures were performed by using a continuous flow helium cryostat (Oxford Instruments, CF935) and a temperature control system (Oxford instruments, ITC 503S). For continuous wave (CW) X-band EPR experiments a cylindrical TE_{011} cavity (Bruker, ER 4119HS) was used. CW EPR spectra of the azurin mutants were recorded at room temperature and at 80 K with a sampling time of 40 ms, a microwave power of 20 mW, a modulation amplitude of 0.1 mT, and a modulation frequency of 100 kHz.

For PELDOR measurements a Flexline probehead housing a dielectric ring resonator (Bruker, ER 4118X-MD5-W1) was used. The second microwave frequency was coupled into the microwave bridge by a commercially available setup (E580-400U) from Bruker. All microwave pulses were amplified via a pulsed travelling wave tube (TWT) amplifier (Applied Systems Engineering, 117X). PELDOR experiments were performed with a standard four-pulse sequence $\pi/2(\nu_{det})$-τ_1-$\pi(\nu_{det})$-$(\tau_1 + t)$-$\pi(\nu_{pump})$-$(\tau_2 - t)$-$\pi(\nu_A)$-τ_2-echo. The frequency of the pump pulse and the magnetic field were adjusted in order to be on resonance with the maximum of the nitroxide spectrum. Then, the frequency of detection pulses was set to be on resonance with a certain part of a Cu^{2+} spectrum. Four different positions of the Cu^{2+} spectrum were probed for each sample, corresponding to frequency offsets of 160, 220, 440, and 560 MHz. The $\pi/2$ and π pulses of the detection sequence had lengths of either 10 and 20 ns or 16 and 32 ns. The amplitudes of these pulses were chosen to optimize the refocused echo. The $\pi/2$ pulse was phase-cycled to eliminate receiver offsets. The length of the pump pulse was optimized to get the maximal inversion of a Hahn echo at the pump frequency. Depending on the frequency offset, the pump pulse had a length of 18-60 ns. The amplitude of the pump pulse was always set to the maximal available value corresponding to an attenuation of 0 dB. The initial τ_1 value was set to 190 ns and was incremented during each experiment 12 times with a step of 8 ns in order to suppress the proton ESEEM. Depending

on the amplitude of the refocused echo, the τ_2 value was adjusted for each sample in the range from 1.0 to 1.5 µs. All PELDOR spectra were recorded at 10 K with a repetition time of 400 µs, a video amplifier bandwidth of 30 MHz and a video amplifier gain of 57 dB. The time increment for the displacement of the pump pulse was set to 4 ns for the azurin mutant $T61R_1$ and to 8 ns for all other azurin mutants. Depending on the sample and on the frequency offset, from 100 up to 700 scans were used to achieve an acceptable signal-to-noise ratio.

3. Biochemical characterization of the azurin samples

Figure S1a shows an SDS-PAGE analysis of azurin at different stages during the purification. Lane 9 shows a sample of spin-labelled azurin $T21R_1$ prior to the PELDOR measurements. All samples showed a single symmetric peak on a Superdex 200 16/60 gelfiltration column (Figure S2). The elution volume is in agreement with monomeric azurin.

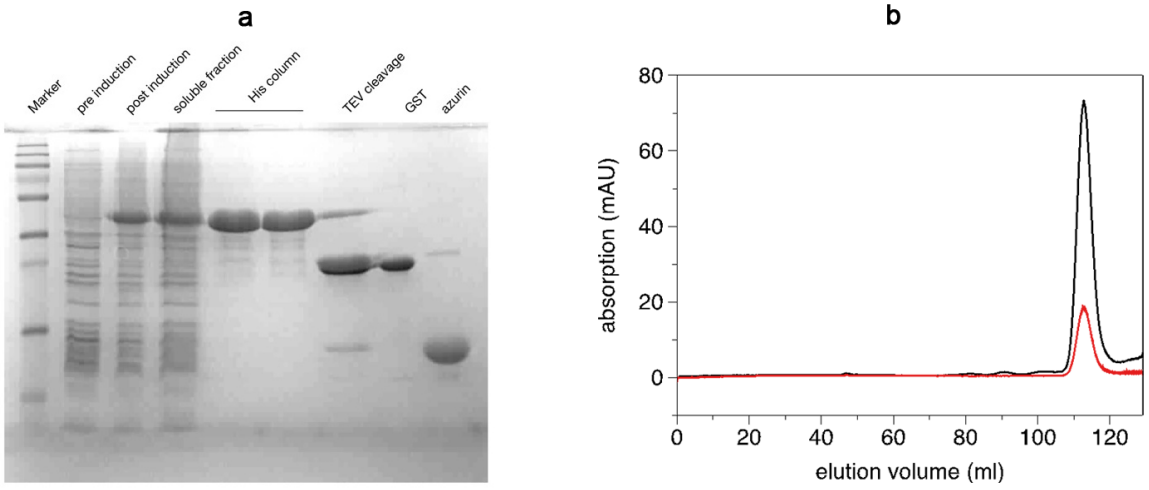


Figure S1. a) SDS-PAGE analysis of azurin at different stages during expression and purification. Lane 1: Marker, Lane 2: Whole cell extract before induction with IPTG, Lane 3: Whole cell extract before harvesting, Lane 4: Soluble fraction of the cell extract, Lanes 5-6: After Ni^{2+}-affinity chromatography and on-column labelling, Lane 7: After cleavage of the GST-tag with TEV protease, Lane 8: Separated GST-tag, Lane 9: Pure spin-labelled azurin after gel filtration chromatography. **b)** Gelfiltration chromatogram of pure spin-labelled azurin mutant $T21R_1$. The black line represents the absorption at 280 nm, the red line - the absorption at 608 nm.

4. EPR characterization of the azurin samples

CW EPR spectra acquired on the azurin mutants are shown in Figure S2. The spectra recorded at room temperature feature the characteristic spectra of protein-bound nitroxides without any traces of free spin labels (Figure S2a). The spectra recorded at 80 K consist of two contributions: the low-field part corresponds to the Cu^{2+} ions, whereas the intense line in the g = 2 region is related to the nitroxide (Figure S2b). One can see that the intensity of the Cu^{2+} signal varies for the different samples, which is attributed to slightly different contents of Cu^{2+} in the azurin samples. Additionally, CW EPR spectra were recorded on the sample of wild time azurin. This sample was prepared via the same procedure as all other azurin mutants, including the mixing with the large excess (~20x) of MTSSL and the further purification via size-exclusion chromatography, to remove all free labels. The room temperature CW EPR spectrum shows no nitroxide signal (Figure S2a), confirming that MTSSL is not binding to wild type azurin.

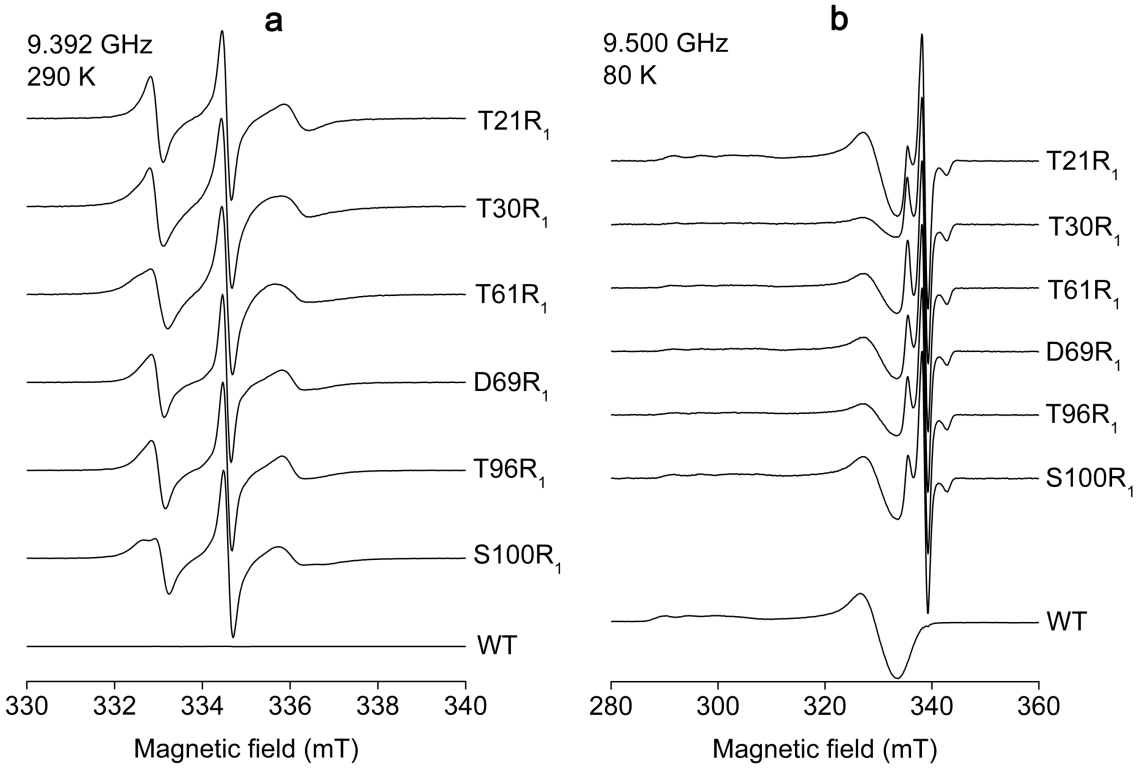


Figure S2. CW spectra of the wild type azurin and its mutants recorded at **a)** room temperature and **b)** 80 K.

5. PELDOR data analysis

The first step of the data analysis was the division of the experimental time traces by a monoexponential decay, to account for intermolecular interactions (Figure S3). Further analysis was done in three different ways.

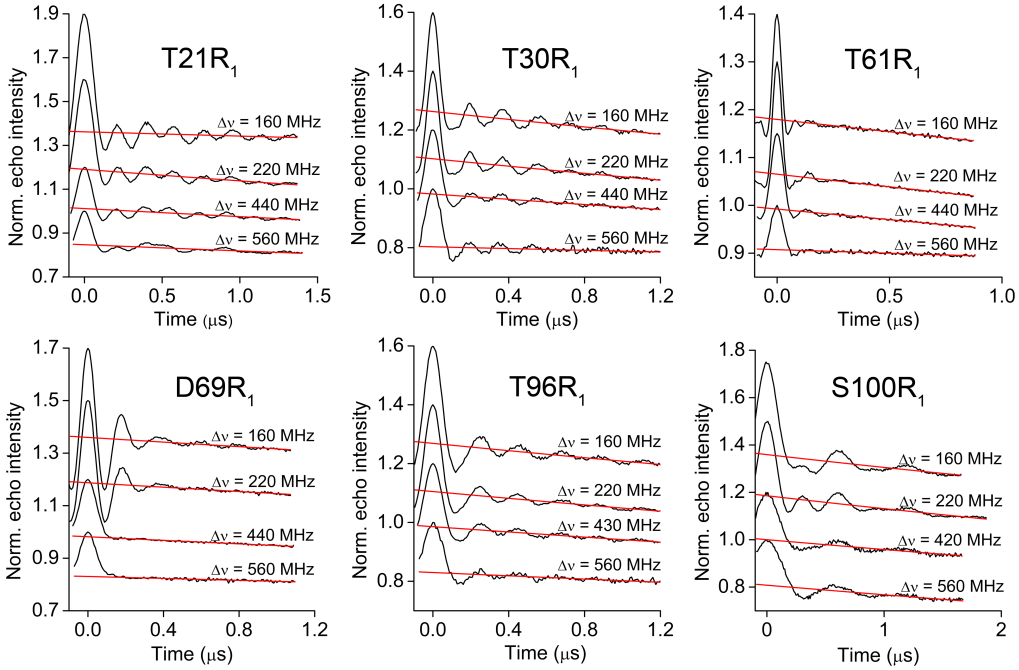


Figure S3. PELDOR time traces of six azurin mutants (black) are overlaid with the background fits (red).

Method 1. The Fourier transform (FFT) of the time traces was calculated in order to obtain the spectra of dipole-dipole interaction. These spectra consist of frequencies which are defined by equation 1,[2]

$$\nu_{dd}(r,\theta) = \frac{D}{r^3}\left(3\cos^2\theta - 1\right), \qquad (1)$$

where r is the distance between the spins in a spin pair, θ is the angle between the spin-spin vector and the external magnetic field, and D = 52.04 MHz/nm^3 when the g-factors of both spins are equal to the electron spin g-factor (2.0023). The frequency corresponding to $\theta = 90°$ ($\nu_\perp$) was determined from the FFTs for each azurin mutant. This frequency was then translated into the most probable distance using equation 1.

Two distinct $\nu_\perp$ frequencies were found in the FFTs of the mutants T21R$_1$, T30R$_1$, T61R$_1$ and T96R$_1$. Therefore, two distances were determined for each of these mutants. As it is mentioned in the main text, these two distances originate from two different conformations of MTSSL.

In order to estimate the error of the obtained distances, the points nearest to $\nu_\perp$ were picked from the FFTs (without zero padding, Figure S4) and converted into two distances using Equation S1 and $\theta = \pi/2$. The interval between these two distances was considered as an error of the method.

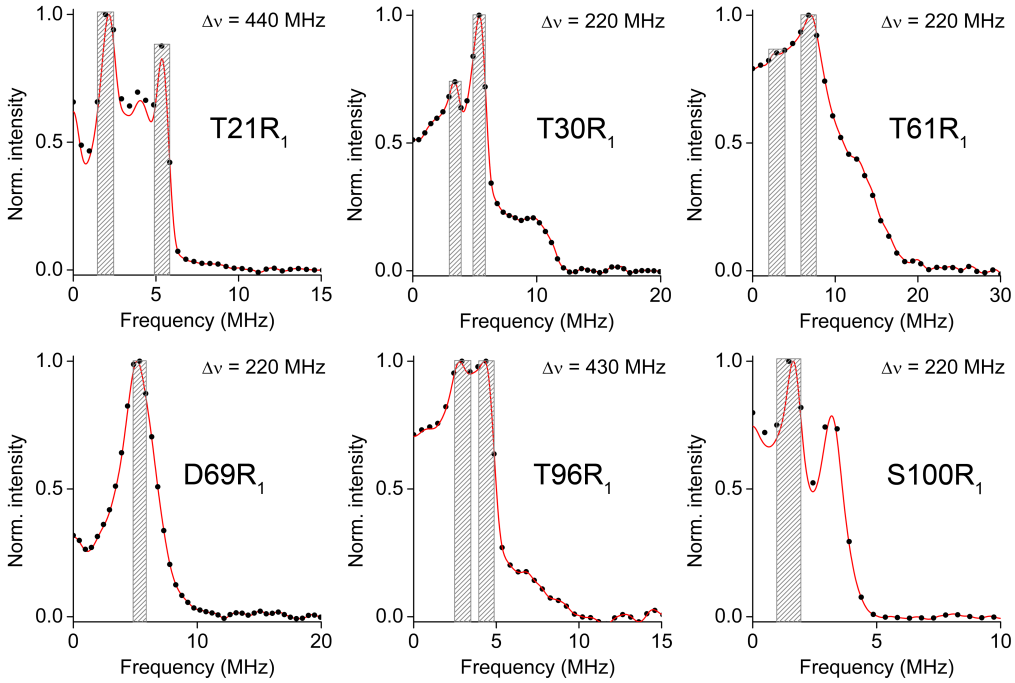


Figure S4. Error intervals for the determination of $\nu_\perp$. FFT of the original and zero-padded PELDOR time traces are shown by dots and red lines, correspondingly. Error bars are shown by rectangles.

Method 2. Four time traces recorded with different frequency offsets were averaged for each mutant and then analyzed with the program DeerAnalysis.[3] The regularization parameter $\alpha = 1$ was used. As can be seen from Figure S5, this method provides relatively good fits to the average time traces and, consequently, to their FFTs. Due to the incomplete orientation averaging, the amplitude of the parallel component of a Pake doublet ($\nu_{||}$, $\theta = 0$ in the Equation S1) is not reproduced for the mutants $T21R_1$ and $D69R_1$. In contrast, a very strong contribution of $\nu_{||}$ into the spectrum of mutant $S100R_1$ is erroneously interpreted by DeerAnalysis as another $\nu_\perp$, corresponding to the distance of 25 Å (see Figure 3c in the main text). In the following, this "ghost" distance peak is excluded from the further analysis.

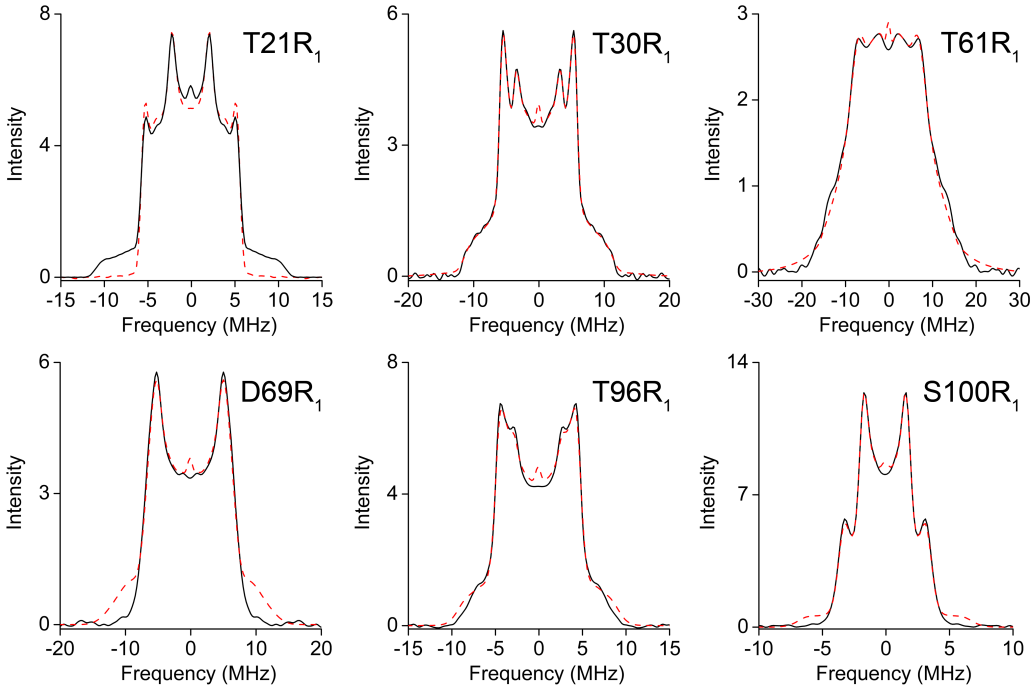


Figure S5. FFT of PELDOR time traces averaged over 4 different offsets are overlaid with their fits as obtained from DeerAnalysis. The corresponding distance distributions are shown in Figure 3c in the main text.

Method 3. The time traces were fitted with the program PeldorFit.[4] In this program, PELDOR time traces are calculated for a simplified model of a spin pair (Figure S6). The geometry of this model is optimized until it provides a fit to all four experimental time traces. To simplify the calculations, the following assumptions are made in the program: 1) both spins are considered as single-point objects; 2) spin-spin distances follow a normal distribution and all angular parameters of the model have a uniform distribution; 3) the correlation between individual geometric parameters of the model is neglected.

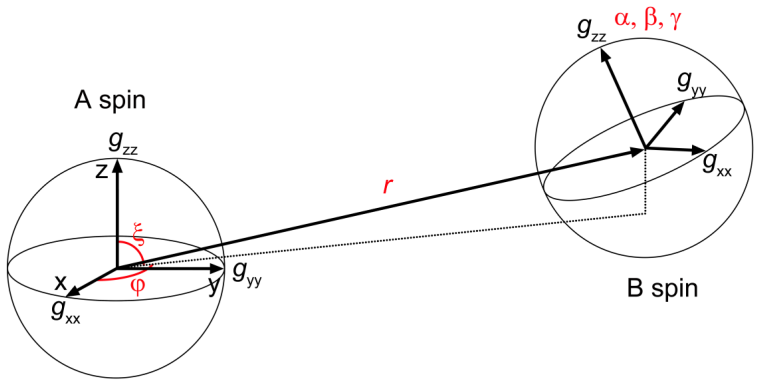


Figure S6. Geometric model of the spin pair. The *g*-tensor of the A spin is supposed to coincide with the coordinate system of the model. The *g*-tensor of the B spin is defined by the Euler angles (α, β, γ) relative to spin A. The Euler angles are defined in accordance with the *z-x'-z"* convention. The inter-spin vector connecting the two spin centers is represented by the spherical coordinates (r, ξ, φ) relative to spin A. All geometric parameters have a distribution. For simplicity, the distance distribution $P(r)$ is approximated by a normal distribution with a mean value μ and a standard deviation σ; α, β, γ, ξ and φ angles are supposed to be equally distributed in the intervals $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$, $\Delta\xi$ and $\Delta\varphi$, respectively.

Table S1. Spectroscopic parameters used for the fitting of the PELDOR time traces.

Parameter	Cu(II)	Nitroxide
g-factor[a]	2.047, 2.047, 2.270	2.0092, 2.0061, 2.0022
g-strain[a]	0.022, 0.022, 0.033	-
A(Cu), MHz[a]	47.0, 47.0, 161.0	-
A(^{14}N), MHz[a]	-	16.2, 16.2, 101.0
Linewidth, MHz[b]	53.3	17.4

[a] The tensors are given in form of tree diagonal components: xx, yy, and zz. [b] The linewidth is given as a peak-to-peak linewidth of the Gaussian shaped line.

In order to correlate the frequency of the mw pulses with the orientations of the spins excited from their EPR spectra, the program requires the spectroscopic parameters of the Cu^{2+} and nitroxide spins. These parameters were obtained from the fitting of the cw EPR spectra of azurin mutants by means of the program EasySpin[5] and are listed in the Table S1. The obtained *g*-tensor and hyperfine tensor of the Cu^{2+} ion are in agreement with the corresponding tensors published previously for wild type azurin.[6] Note that a slight rhombicity of the Cu^{2+} *g*-tensor was found at W-band frequencies,[6a] however at X-band frequencies it is not resolved which allows one to consider the *g*-tensor of Cu^{2+} as being axial. The hyperfine tensor of the Cu^{2+} ion has also an axial symmetry and its z-component is nearly collinear to the z-component of the *g*-tensor.[6c] This symmetry considerations allow one to reduce the number of optimization parameters by setting φ angle and its distribution width to 0 for the geometric model.

The model used in Ref. [4] allows only a unimodal distribution of spin-spin distances. To allow for a bimodal distribution of spin-spin distances, all geometric parameters were duplicated for the fitting of the time traces of mutants $T21R_1$, $T30R_1$, $T61R_1$ and $T96R_1$. Hence, the distance distribution can be approximated by a sum of two Gaussian functions. The ratio between both conformations can also be varied. In order to reduce the number of fitting parameters for these four mutants, the distribution width of the angular parameters was set to 0. This is a valid approximation for these mutants, because the flexibility of the spin label is significantly reduced in these samples, which is reflected by the small widths of the distance distributions obtained in all three methods.

The obtained fits are shown in Figure 3a in the main text. Geometric parameters of the Cu^{2+}-MTSSL spin pair used as fitting parameters are listed in Table S2 together with their optimized values. Note that a good fit to the PELDOR time traces of the mutants $T21R_1$, $T30R_1$, $T61R_1$ and $T96R_1$ was only achieved when a bimodal distribution of spin-spin distances was employed.

Table S2. Optimized parameters of the geometric models.

Mutant	μ, σ (nm)	ξ, $\Delta\xi$ (°)	φ, $\Delta\varphi$ (°)	α, $\Delta\alpha$ (°)	β, $\Delta\beta$ (°)	γ, $\Delta\gamma$ (°)	η[a]
$T21R_1$	2.15, 0.03 (51%) 2.84, 0.08 (49%)	36, - 71, -	-	116, - 178, -	100, - 20, -	3, - 10, -	0.98 (160 MHz) 0.82 (220 MHz) 0.92 (440 MHz) 0.90 (560 MHz)
$T30R_1$	2.12, 0.04 (84%) 2.45, 0.06 (16%)	73, - 50, -	-	93, - 109, -	40, - 43, -	169, - 4, -	0.94 (160 MHz) 0.87 (220 MHz) 0.98 (440 MHz) 1.00 (560 MHz)
$T61R_1$	1.90, 0.10 (86%) 2.38, 0.32 (14%)	77, - 78, -	-	13, - 91, -	29, - 39, -	61, - 3, -	0.47 (160 MHz) 0.53 (220 MHz) 0.66 (440 MHz) 0.47 (560 MHz)
$D69R_1$	2.06, 0.13	14, 28	-	4, 6	37, 9	125, 61	0.90 (160 MHz) 0.86 (220 MHz) 0.94 (440 MHz) 0.94 (560 MHz)
$T96R_1$	2.27, 0.04 (68%) 2.55, 0.13 (32%)	66, - 88, -	-	100, - 91, -	42, - 64, -	71, - 148, -	0.97 (160 MHz) 0.95 (220 MHz) 0.99 (430 MHz) 0.96 (560 MHz)
$S100R_1$	3.08, 0.09	83, 12	-	40, 40	73, 7	56, 38	0.98 (160 MHz) 0.92 (220 MHz) 0.91 (420 MHz) 0.86 (560 MHz)

[a] η is a phenomenological factor which takes into account factors affecting the modulation depth parameter: non-ideal pulses and incomplete spin labelling etc. It can take values in the range of $0 \leq \eta \leq 1$. During the fitting the modulation depth of PELDOR time traces are multiplied by this value.[4]

To estimate the precision of the obtained distances, the root-mean-square deviation (RMSD) between the experimental time traces and their fits was calculated for each mutant as a function of the mean distance μ and its standard deviation σ. During this calculation μ and σ were varied, whereas all other geometric parameters of the model were set to their optimized values. Note that it would be more robust to optimize all other parameters for each pair of μ and σ, but this task is too time consuming. Two RMSD surfaces were computed for the mutants $T21R_1$, $T30R_1$, $T61R_1$ and $T96R_1$ corresponding to two different distances as determined above. All calculated RMSD surfaces have a single minimum whose coordinates are in agreement with the distances obtained after the fitting (Figure S7). For the samples with two distances one can see that one distance peak having a higher probability is well resolved, whereas the second distance peak is usually less exact. In order to provide some quantitative estimation of μ and σ how errors, a range in which 90% of the minimal RMSD is reached were determined for each RMSD surface (Table S3). These ranges can be considered as a rough estimate of the confidence intervals for the corresponding parameters.

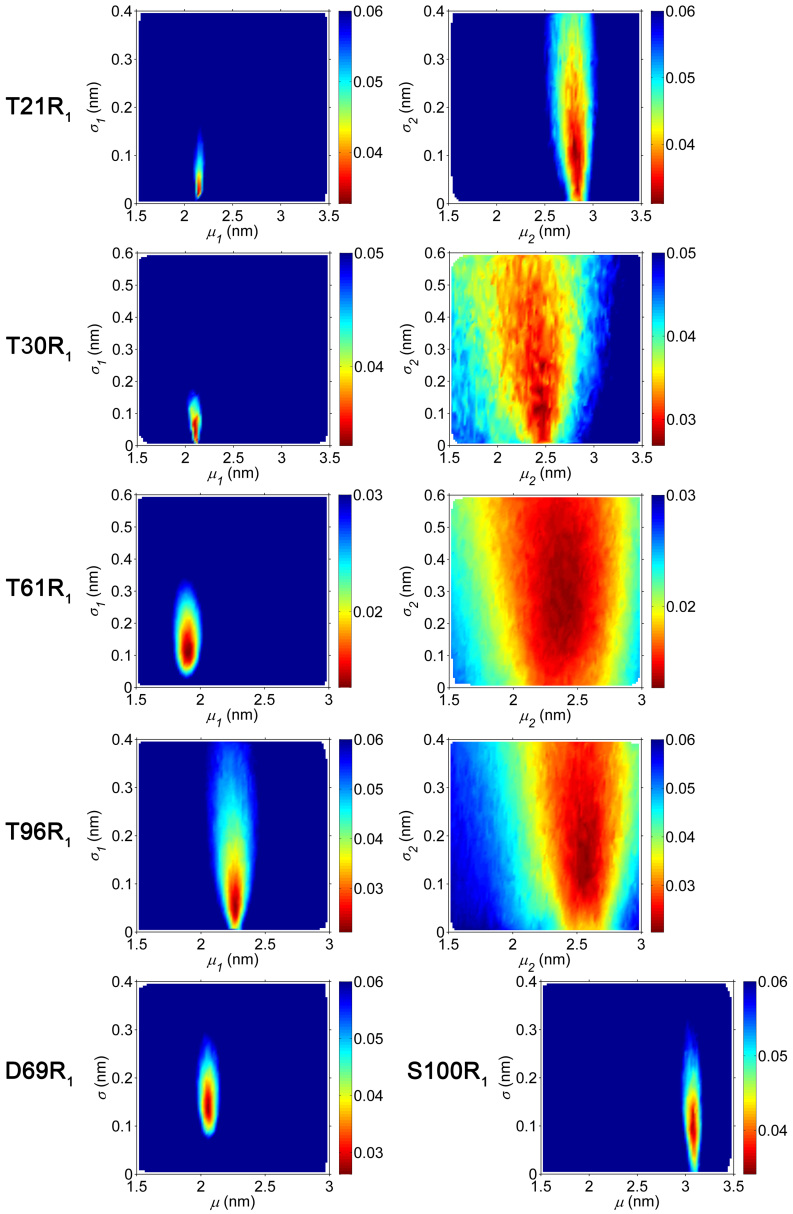


Figure S7. RMSD surfaces for the distances determined from the fitting of PELDOR data.

Table S3. The distance parameters and their errors obtained by PeldorFit.

Mutant	μ (nm) [a]	σ (nm) [a]
$T21R_1$	2.15 (0.01) 2.84 (0.09)	0.03 (0.01) 0.08 (0.13)
$T30R_1$	2.12 (0.03) 2.45 (0.23)	0.04 (0.07) 0.06 (0.39)
$T61R_1$	1.89 (0.05) 2.38 (0.34)	0.10 (0.05) 0.32 (0.46)
$D69R_1$	2.06 (0.01)	0.13 (0.04)
$T96R_1$	2.27 (0.02) 2.55 (0.13)	0.04 (0.02) 0.13 (0.18)
$S100R_1$	3.08 (0.02)	0.09 (0.06)

[a] The values in brackets determine the range in which 90% of the minimal RMSD is reached.

6. Trilateration procedures

Main procedure. The location of the Cu^{2+} ion in the azurin structure was calculated by means of the mtsslTrilaterate program.[7] This program uses the average coordinates of each spin label and the corresponding Cu^{2+}-MTSSL distances as an input data. The distances were taken from the PELDOR data analysis described above. To establish the MTSSL coordinates, models of MTSSL were generated within the molecular coordinate system of azurin (PDB 1E65) by means of the simulation program mtsslWizard[8] (Figure S8). Due to the flexibility of MTSSL, each model consists of many conformers of the spin label. Therefore, the average coordinates are calculated for each spin label while the generated models of MTSSL are imported into mtsslTrilaterate.

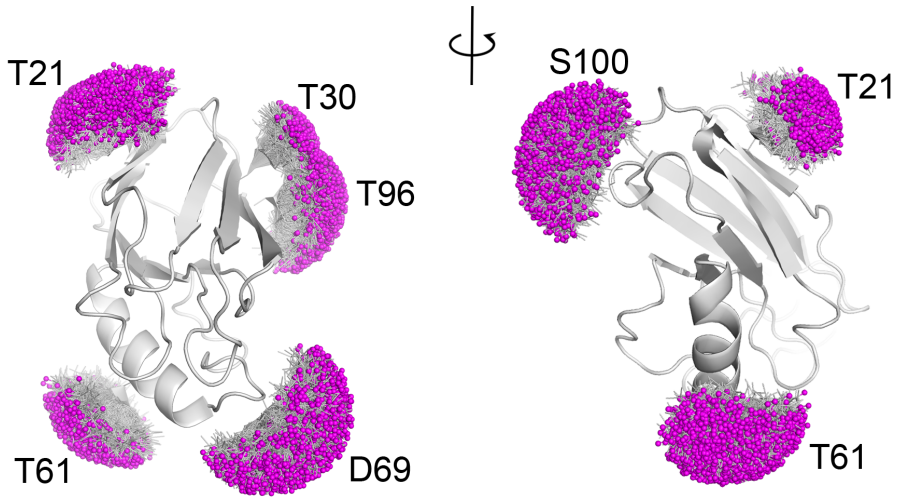


Figure S8. Models of MTSSL attached to the azurin surface as obtained from the mtsslWizard program. The *apo* structure of azurin (PDB 1E65) is shown as a cartoon. MTSSL conformers are shown by sticks with spheres corresponding to the oxygen atoms.

As soon as all these input data is plugged into mtsslTrilaterate, the trilateration problem is solved using a nonlinear least square routine.[7] After the calculation is finished, the program outputs the most probable Cu^{2+} coordinates and their standard errors. The position in the azurin structure which corresponds to these coordinates is than visualized using the PyMOL program (http://www.pymol.org/). The calculated Cu^{2+} position is depicted as an ellipsoid with the center set to the most probable coordinates and semi-principal axes set to the doubled standard errors (2σ, 95% confidence level).

Influence of the precision of distance constraints on trilateration. Trilateration of the Cu^{2+} ion was performed using the three different sets of distance constrains as obtained from the FFT approach, DeerAnalysis and PeldorFit (see Section 5). All three sets are given in Table 1 in the main text. The corresponding solutions are shown within the azurin structure and

compared to the crystallographic position of Cu^{2+} in Figure S9. The solution corresponding to the PeldorFit-derived distances has the smallest shift from the crystallographic position, 2.6 Å. The similar trilateration result is obtained for the distances derived from DeerAnalysis (2.7 Å). In contrast, a significantly bigger error is obtained for the FFT approach (4.2 Å). This is expected, since the FFT approach is less accurate in comparison to other two methods.

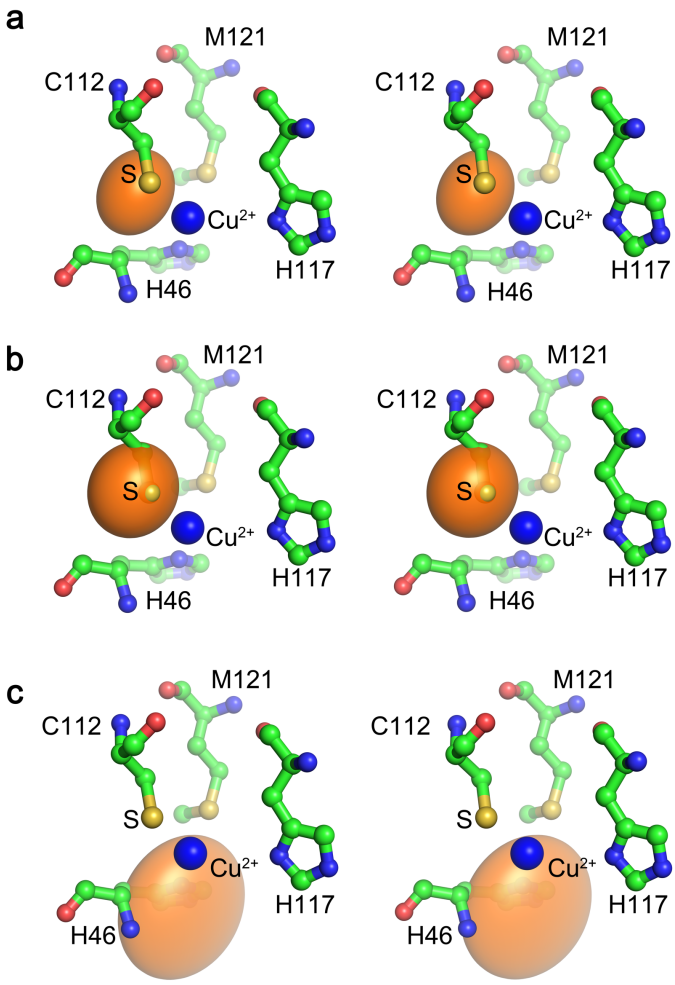


Figure S9. Trilateration of the Cu^{2+} ion in azurin for distance constraints derived from **a)** PeldorFit, **b)** DeerAnalysis, and **c)** the FFT approach (stereo pairs). The calculated 2σ area of the Cu^{2+} ion location is depicted by an orange ellipsoid, and the corresponding crystallographic position is shown by a blue sphere. The amino acid site chains comprising the binding site are shown as ball-and-stick models.

Influence of the number of distance constraints on trilateration. The impact of the number of constraints on the trilateration result was studied by reducing the number of constraints from 6 down to 5 and 4 (Figure S10). Firstly, the trilateration was attempted with 15 different combinations of 4 distance constraints, which is the minimum number of constraints for the algorithm to obtain a unique trilateration solution. The corresponding solutions place the Cu^{2+} ion away from its crystallographic position by 1.4 to 13.9 Å, resulting in an average deviation

of 4.4 Å. In the next step, 6 combinations of 5 distance constraints were used for trilateration. For this input data the difference between the calculated and crystallographic Cu^{2+} positions vary between 2.7 and 3.2 Å, yielding an error of 2.9 Å in average. The last case of 6 constraints corresponds to the difference of 2.6 Å.

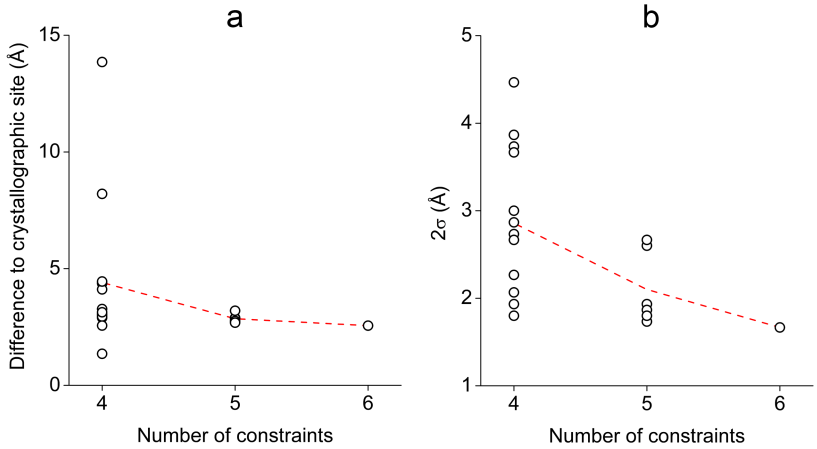


Figure S10. a) The difference of the calculated Cu^{2+} position to the corresponding crystallographic position as a function of the number of distance constraints used for trilateration. **b)** The average 2σ-error of the calculated Cu^{2+} coordinates as a function of the number of distance constraints used for trilateration.

Influence of the protein structure on trilateration. In addition to the *apo* structure of azurin, a homology model of azurin was tested as a reference structure. The homology model was derived from the crystal structure of the related blue copper protein auracyanin (PDB 2AAN, 30% identical amino acids) using the SWISS-MODEL server.[9] Superimposed onto the *apo* structure, the homology model shows an RMSD of 1.26 Å based on 103 C_α atoms (Figure S11a). Later, the models of MTSSL attached to the positions 21, 20, 61, 69, 96, and 100 of the homology model were created by means of mtsslWizard. Since *in silico* spin labeling is rather sensitive towards the surrounding of the labeled site, it is worth noting that the obtained distributions of MTSSL are different for the *apo* and homology models. Using the average coordinates of the newly generated MTSSL models and the distances from PeldorFit, the position of the Cu^{2+} ion was calculated. The obtained position is again close to the expected binding site of the Cu^{2+} ion (Figure S11b). Note that due to the prominent discrepancies of the *apo* and homology models, such a good accuracy of the obtained Cu^{2+} position was not expected. However, this result should not be over-interpreted, because it could be simply due to the compensation of errors stemming from the individual input parameters (e.g. average coordinates of different spin labels).

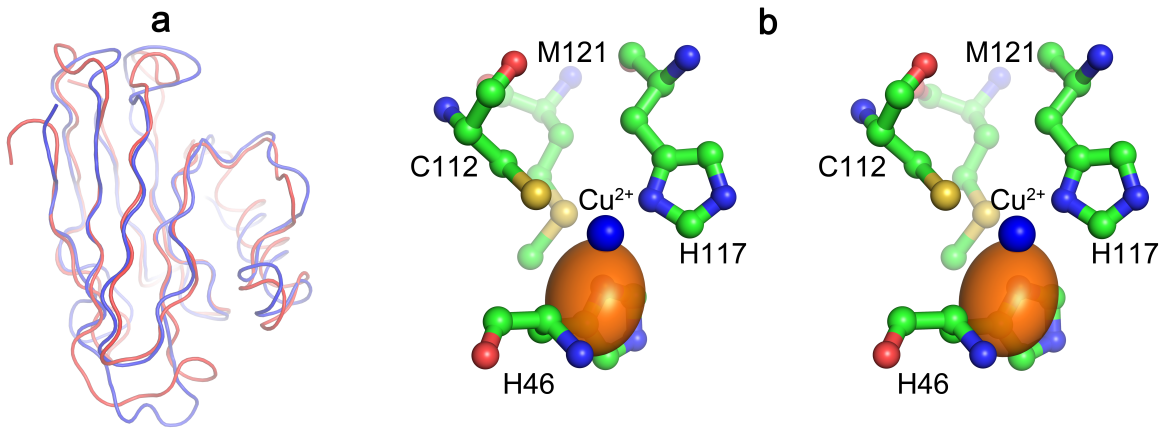


Figure S11. a) The structure of the homology model (blue) of azurin is overlaid with the *apo* structure of azurin (red). **b)** The trilateration result for the homology model of azurin, which was used as a reference structure to generate the models of the spin labels. The calculated 2σ confidence area of the Cu^{2+} ion location is depicted by the ellipsoid. The amino acid site chains comprising the binding site are shown as ball-and-stick models.

References

[1] H. Liu, J. Naismith, *BMC Biotechnology* **2008**, 8:91.

[2] O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1-53.

[3] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473-498.

[4] D. Abdullin, G. Hagelueken, R. I. Hunter, G. M. Smith, O. Schiemann, *Mol. Phys.*, DOI: 10.1080/00268976.2014.960494.

[5] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42-55.

[6] a) J. W. A. Coremans, O. G. Poluektov, E. J. J. Groenen, G. W. Canters, H. Nar, A. Messerschmidt, *J. Am. Chem. Soc.*, **1994**, *116*, 3097-3101; b) J. E. Roberts, J. F. Cline, V. Lum, H. B. Gray, H. Freeman, J. Peisach, B. Reinhammar, B. M. Hoffman, *J. Am. Chem. Soc.*, **1984**, *106*, 5324-5330; c) W. E. Antholine, P. M. Hanna, D. R. McMillin, *Biophys. J.* **1993**, *64*, 267-272.

[7] G. Hagelueken, D. Abdullin, R. Ward, O. Schiemann, *Mol. Phys.* **2013**, *111*, 2757-2766.

[8] G. Hagelueken, R. Ward, J. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377-391.

[9] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics* **2006**, *22*, 195-201.

[P5] Determination of nitroxide spin label conformations *via* PELDOR and X-ray crystallography

Reprinted with permission from

D. Abdullin, G. Hagelueken, O. Schiemann, "Determination of Nitroxide Spin Label Conformations via PELDOR and X-ray Crystallography", *Phys. Chem. Chem. Phys.* **2016***, 18,* 10428-10437.

DOI: 10.1039/c6cp01307d

Own contribution to the manuscript:

- Calculation of MTSSL conformers from PELDOR data
- Comparison of the PELDOR and X-ray results
- Writing the manuscript

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 10428

Determination of nitroxide spin label conformations *via* PELDOR and X-ray crystallography†

D. Abdullin,‡ G. Hagelueken‡ and O. Schiemann*

Received 25th February 2016,
Accepted 11th March 2016

DOI: 10.1039/c6cp01307d

www.rsc.org/pccp

Pulsed electron–electron double resonance (PELDOR or DEER) in combination with site-directed spin labelling has emerged as an important method for measuring nanometer distance constraints that are used to obtain coarse-grained structures of biomolecules or to follow their conformational changes. Translating measured spin–spin distances between spin labels into structural information requires taking the conformational flexibility of spin label side chains into account. Here, we present an analysis of orientation selective PELDOR data recorded on six singly MTSSL-labelled azurin mutants. The analysis yielded conformational MTSSL ensembles, which are considerably narrower than those predicted using *in silico* spin labeling methods but match well with spin label conformations found in the corresponding crystal structures. The possible reasons and consequences for predicting spin label conformers in the fold of biomolecules are discussed.

Introduction

Living cells are teeming with macromolecules, whose complex three-dimensional structures, interactions and dynamics are of crucial importance for all cellular processes and ultimately the cell's survival. Although structural biologists have studied the structures of macromolecules for many decades, new methods are constantly needed to investigate the increasingly complex questions that arise from different fields of biology. Pulsed electron–electron double resonance (PELDOR, also known as double electron–electron resonance or DEER)[1] spectroscopy has developed into a method for obtaining long-range distance constraints for macromolecules in solution. The experiment can be performed in frozen[2] solutions or at room temperature[3] and within cells.[4] PELDOR allows measuring distances between two or more spin centres, *e.g.* paramagnetic metal ions,[5] organic cofactors,[6] or spin labels,[2,7] in the range of 15–150 Å.[8] The distances obtained from PELDOR experiments can be used to study conformational changes,[9] to validate structural models,[10] to localise metal ions[11] or to produce coarse-grained structures.[12] In the case of proteins, the most widely used spin label is a methanethiosulfonate spin label (MTSSL),[13] which binds covalently to cysteines, forming the so-called R1 side chain (Fig. 1). However, the flexibility of the R1 side chain poses an important challenge in the interpretation of experimental PELDOR data: the PELDOR experiment provides the distance between the spin centres, which are localised on the NO groups of the R1 side chains, while it is the distance between the C_α atoms, which is of actual interest to the structural biologist. Due to the intrinsic flexibility and ~7 Å length of the R1 side chain these two distances are often not easy to correlate. Thus, in order to relate the experimental distance to the underlying protein structure, one has to know, or at least estimate the conformation of the R1 side chain in its macromolecular environment, – a task that is far from trivial.

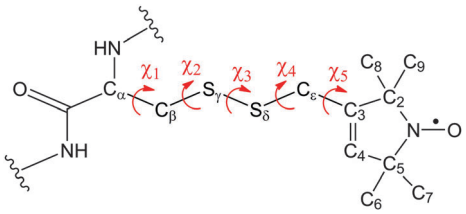


Fig. 1 Structure of the R1 side chain. The spin is localized to more than 95% between the N and O atoms. The five dihedral angles χ_1–χ_5 corresponding to the rotatable bonds of R1 are shown by red arrows.

In some studies, this task was approached experimentally. For example, Abé *et al.*[14] and Endeward *et al.*[15] were able to obtain orientation selective PELDOR time traces and to derive the most probable conformations of relatively tightly packed R1 side chains. In addition, continuous wave EPR-based mobility studies in combination with molecular dynamics (MD) simulations were used to qualitatively study R1 conformations.[16] Also crystal

Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany. E-mail: schiemann@pc.uni-bonn.de

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp01307d

‡ These authors contributed equally to this work.

structures of several R1-labelled proteins (*e.g.* T4 lysozyme,[17] Wza,[18] BtuB,[19] Spa15,[20] CylR2,[21] GB1,[22] cytochrome P450cam[23] and azurin[24]) have been obtained revealing that the label is often involved in many interactions with the protein surface and that the nitroxide group itself can interact with polar groups of the protein or form hydrogen bonds with solvent molecules. Whereas the latter interactions have also been observed in solution by EPR spectroscopy,[25] it is unclear whether or in how far the spin label conformations observed in the crystal are also relevant for the frozen solution.

The lack of a straightforward experimental method for obtaining spin label conformations in solution led to the use of computational methods. For example, Schiemann *et al.* used MD simulations to interpret PELDOR distances for the spin-labeled DNA[26] and later Sale *et al.*[27] showed that MD-derived models of spin-labeled proteins enable the interpretation of mean PELDOR distances with an accuracy of $\sim$3 Å. Since then, several improvements to the force fields, solvent models and time scales of the calculated trajectories have been made, but no significant change in the accuracy of the distance predictions has been reported so far.[28] As an alternative to the time consuming all-atom MD simulations, several simplified *in silico* spin labeling methods have been proposed. These methods compute R1 conformations based on an accessible volume approach,[29] a weighted rotamer library approach[30,31] or an MD approach with dummy spin labels.[32] Although these methods differ markedly in complexity and sophistication, an average mean distance prediction error of $\sim$2–3 Å was found for all of them. Beyond the error of the mean distance, the shapes of the computed distributions often do not fit well to the experimentally determined shapes, even if high quality PELDOR data were obtained.[32,33] Consequently, although the current precision of *in silico* labelling programs has been shown to be good enough for coarse-grained modelling,[12] many distance constraints are still required for reducing the total error. In addition, many important biological processes are governed by small-scale conformational changes at specific points of a structure. Often, only very few possible labelling sites exist in such cases, making it much more important to predict the spin label conformation with the best possible precision. It is thus of great interest to reduce the error of *in silico* spin labelling predictions and a detailed investigation of the interactions of the spin label with its environment is therefore essential to improve the prediction methods.

In principle, PELDOR spectroscopy offers an elegant possibility of determining the orientation of the spin labels *in situ*, by orientation selective excitation of spin centres. Provided that the contributions of the individual spin centres to the orientation selectivity can be disentangled, the conformations of the spin labels can be determined. Here, this problem is approached by analysing an orientation selective PELDOR data set[11] of Cu^{2+}–R1 distance measurements on six R1-labelled azurin mutants. The Cu^{2+} ion is immobilised in the azurin structure and its magnetic tensors are well-known.[11,34] This makes it possible to determine the positions of the NO group of the R1 side chain relative to the Cu^{2+} ion from the orientation selectivity and further to construct experimental models of the R1 side chains in frozen solution. These models are then compared to a previously determined crystal structure of azurin T21R1[24] and two newly determined crystal structures of azurin T30R1. The findings reveal insights into the behaviour of the R1 side chain in the crystal lattice and frozen solution, which may help improve the existing *in silico* spin labelling algorithms.

Experimental

The synthesis and spin labelling of the six azurin mutants (T21R1, T30R1, T61R1, D69R1, T96R1, S100R1) as well as the corresponding PELDOR experiments are described elsewhere.[11] Briefly, the PELDOR measurements were carried out using an X-band EPR spectrometer (Bruker ELEXSYS E580). Four PELDOR time traces were acquired for each of the azurin mutants. The frequency of the pump pulse was set in resonance with the maximum of the nitroxide spectrum, and the frequency of the detection pulses was successively adjusted to four different positions on the Cu^{2+} spectrum (see Table S1 for details, ESI†). The difference between the frequencies of the pump and detection pulses, the so-called frequency offset $\Delta\nu$, was either 160, 220, 440 or 560 MHz. All obtained PELDOR signals were divided by an exponential decay, to account for intermolecular interactions.

Crystallization, data collection, structure solution and refinement of azurin T21R1 have been previously described.[24] Crystals of T30R1 were found in condition 2.21 of the JCSG+ screen (Molecular Dimensions). The conditions were optimised to produce larger crystals (1.50–2.25 M sodium malonate 5 M pH 7.4). Two different shapes of crystals (I and II) were observed. Crystals from both shapes were cryoprotected with 35% glycerol and flash-cooled in liquid nitrogen. Diffraction datasets were collected at beamline BL14.3 of the BESSYII synchrotron (Berlin, Germany). Crystals of shape I were indexed, integrated and scaled in space group *C*2 using IMOSFLM[35] (Table S2, ESI†). The structure was solved using PHASER[36] (search model 1E67[37]) and contains four monomers per asymmetric unit. The structure was refined using PHENIX.REFINE[38] until R/R_{free}-factors converged at 0.194/0.240. The geometric quality of the model was checked using MOLPROBITY.[39] Crystals of shape II were initially indexed as $I2_12_12_1$ and PHASER (search model 1E67) found 8 monomers in the asymmetric unit. During refinement with PHENIX.REFINE the *R*-factor stalled at $\sim$0.34, which was suspiciously high for a 1.95 Å resolution structure. The dataset was therefore reprocessed in space group *C*2. The structure then contained 16 crystallographically independent monomers and could be refined to *R*-factors of 0.212/0.271. The geometry of the model was evaluated using MOLPROBITY (Table S2, ESI†).

Computational methods

The algorithm used to derive the conformations of the R1 side chain from the corresponding PELDOR data sets can be divided into three parts:

Generation of the trial R1 conformers

The calculations start with the generation of an initial ensemble of trial R1 conformers using either mtsslWizard[29] or MMM[30] and the crystal structure of native azurin (PDB 1E67) as a protein model. In each case, 10^4 or 200 different trial conformers of R1 were created using mtsslWizard and MMM, respectively. To account for the mobility and possible changes of the protein side chains around the spin label, the "loose mode" of mtsslWizard was used. The MMM R1 conformers were calculated using the rotamer library "R1A_298K_UFF_216_r1_CASD". For the following calculations all rotamers with non-zero weight were used and their weights were discarded.

Simulation of the PELDOR time traces for trial R1 conformers

In the next step, PELDOR time traces were simulated for each of the trial R1 conformers and the Cu^{2+} ion of azurin using the same frequencies and lengths of the microwave pulses as in the actual experiments (Table S1, ESI†). For these simulations, the relative orientation of each nitroxide and the Cu^{2+} spin was described by six parameters: the length of the distance vector, two polar angles defining the direction of the distance vector with respect to the *g*-tensor of the Cu^{2+} spin, and three Euler angles defining the orientation of the nitroxide *g*-tensor relative to the *g*-tensor of the Cu^{2+} spin. The orientation of the Cu^{2+} *g*-tensor in azurin was taken from the literature,[34] whereas the orientation of the nitroxide *g*-tensor relative to the Cu^{2+} *g*-tensor was calculated from the geometry of each trial conformer.[40] PELDOR simulations were done according to the procedure described earlier.[11,41] All simulated PELDOR signals were averaged over 10^6 random orientations of the magnetic field. Additionally, the delocalisation of 60% of the copper spin density onto the sulphur atom of the residue C112 was taken into account when calculating dipolar frequencies.[5]

Selection of the best R1 conformers

Next, the subset of the mtsslWizard- or MMM-generated trial conformers, which provides the best fit to the experimental PELDOR signals, was determined. In order to solve this optimisation problem, a genetic algorithm was employed. This algorithm has been shown to be very efficient when dealing with a large number of optimisation parameters.[41,42] As mentioned above, the optimisation target is an ensemble of R1 conformers that produce the best possible fit between simulated and experimental PELDOR data, where each conformer in a particular ensemble represents one optimisation parameter. For simplicity, the size of the target ensemble was set to a constant value N_g during optimisation. The genetic algorithm can then pick any of the trial conformers and assign it to one of the N_g optimisation parameters. The possibility of varying occupancies for the N_g conformers was taken into account by simply allowing for picking a certain conformer multiple times.

In the terminology of the genetic algorithm, each of the N_g optimisation parameters and hence each selected trial conformer is a "gene". A set of N_g genes determines one complete set of optimisation parameters and is a "chromosome". A set of chromosomes with a fixed size N_c forms a "generation". In order to decide which chromosomes of the generation best fulfil the optimisation criteria, the generation undergoes a "scoring" process. During scoring, a root-mean-square deviation (RMSD) between the experimental and simulated PELDOR signals was calculated for each chromosome. The simulated PELDOR signals were obtained *via* summation of the precomputed PELDOR signals of the selected trial R1 conformers defined by the genes of each chromosome. The RMSD value was calculated as a sum of RMSDs for four frequency offsets. This ensures a global fit of all four time traces that were recorded for each azurin mutant.

In the genetic algorithm, each iteration cycle includes the following steps: (1) creation of a new generation, (2) scoring, (3) determination of the best chromosome. Except for the first generation, where all genes are chosen randomly, every new generation is created from the previous one *via* three evolutionary processes – selection of $N_c/2$ pairs of the "parent" chromosomes, crossover of their genes with the probability p_c, and mutation of their genes with probability p_m. Each pair of the parent chromosomes is selected by randomly picking two pairs of chromosomes from the previous generation and choosing from each pair the chromosome with the lower RMSD (tournament selection). The genetic algorithm runs this way until the number of iterations (generations) reaches the maximum N_{max}, which is specified beforehand.

One notable feature of the genetic algorithm is that its performance depends on the values of its intrinsic parameters, such as the size of the generation (N_c), the crossover and mutation probabilities (p_c and p_m), and the maximal number of generations (N_{max}). Since the optimal values of these parameters may differ for different problems, they have to be adjusted for each individual problem, mostly empirically. Such an adjustment was done here by running the genetic algorithm for different values of each of the parameters and setting other parameters to a constant value (Fig. S1, ESI†). A good performance of the algorithm was achieved with N_c = 500, p_c = 0.5, p_m = 0.01, and N_{max} = 10^4. Additionally, the number of fitting parameters or genes per chromosome (N_g) was tuned in a similar way. From 1 up to 100 genes per chromosome were used to run the genetic algorithm while all parameters of the genetic algorithm were set to their optimised values. Our tests revealed that at least 20 genes are required to obtain good fits to the PELDOR signals (Fig. 2).

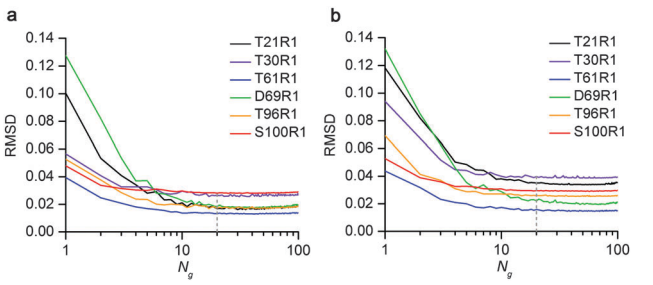


Fig. 2 RMSD between the simulated and experimental PELDOR time traces of the six azurin mutants as a function of the number of R1 conformers used for the simulation: (a) mtsslWizard conformers, (b) MMM conformers. Other parameters of the genetic algorithm are N_c = 500, p_m = 0.01, p_c = 0.5, N_{max} = 10^4. The dotted line indicates the value of 20 used for N_g.

Results and discussion

Identification of preferred R1 conformations using orientation selective PELDOR data

For each of the six azurin mutants, a data set of four orientation selective PELDOR time traces was available from previous work.[11] As mentioned above, the high data quality and the fact that the azurin mutants contain only one flexible spin label should make it possible to determine the location of the NO group within the fold of azurin. The time traces of the mutants T21R1 and T30R1 are reproduced in Fig. 3 (see Fig. S2 for the other mutants, ESI†). The program PeldorFit was previously used to extract the Cu^{2+}–nitroxide distance distributions from these sets of orientation selective time traces. The distance distributions of all mutants were found to consist of one or two narrow peaks, suggesting the presence of preferred R1 conformations in the samples (Fig. 4). In addition to the distances, the PeldorFit analysis yielded the relative orientations of Cu^{2+} and nitroxide spins. Although these orientations encode information about the possible conformations of the R1 side chain, there is no easy way to extract this information. Instead, we decided to search for the underlying R1 conformations by using structural models of R1 from two *in silico* spin labelling programs, mtsslWizard and MMM. The program mtsslWizard implements an accessible volume approach and generates random, non-weighted models of the R1 side chain discarding only those conformations that clash into the protein or the label itself. In contrast, the MMM-derived models are based on a rotamer library, which consists of energetically favourable conformations of a free R1 side chain. MMM also takes some attractive non-bonding interactions such as the C_α–S_δ interaction[43] into account. In order to estimate the accuracy of the mtsslWizard- and MMM-generated R1 ensembles, their Cu^{2+}–NO distance distributions were calculated and compared with the experimental PELDOR distances (Fig. 4). On average, the mean distances of the calculated and experimental distributions deviate by only 1.5 Å. However, all calculated distributions are much wider than the experimentally determined ones. For example, the case of T21R1 demonstrates that a good estimate of the experimental mean distance can be achieved even if the shapes of the predicted and experimental distance distributions deviate significantly from each other. Obviously, high-quality experimental data are necessary for such a shape comparison.[33] The discrepancies are especially prominent for mutants T21R1, T30R1 and T96R1, for which narrow bimodal distance distributions have been experimentally determined. Interestingly, all these three mutants have the spin label located in solvent exposed β-sheets,[44] which have a tighter packing of the side chains as compared, for example, to solvent exposed residues on α-helices. As a consequence of the differences between the experimental and calculated Cu^{2+}–NO distance distributions, the PELDOR time traces that are simulated for the mtsslWizard- and MMM-generated R1 ensembles do not agree with the experimental time traces (Fig. 3a and b). The faster modulation damping in the simulated time traces as compared to the experimental time traces indicates that the *in silico* models overestimate the conformational diversity of the spin labels. Therefore, it should be possible to find sub-ensembles of the initial ensembles, which provide a better fit to the PELDOR data and are thus a better model of the R1 side chains.

A genetic algorithm was used to solve this optimisation problem. Briefly, the algorithm iterates over various combinations of trial conformers until a combination of 20 trial conformers is found that provides the best fit to the four experimental

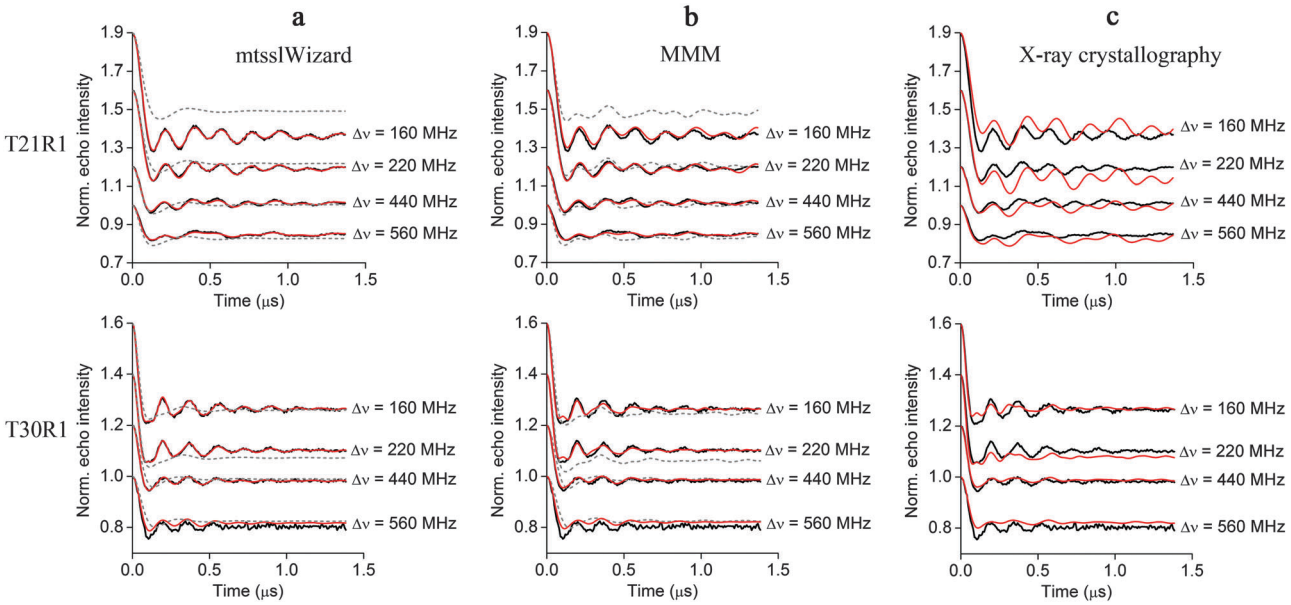


Fig. 3 Simulation of the experimental PELDOR time traces (black) for azurin mutants T21R1 and T30R1 (a) for the initial (grey) and optimised (red) mtsslWizard-generated R1 conformers, (b) for the initial (grey) and optimised (red) MMM-generated R1 conformers, and (c) for the X-ray R1 conformers (red).

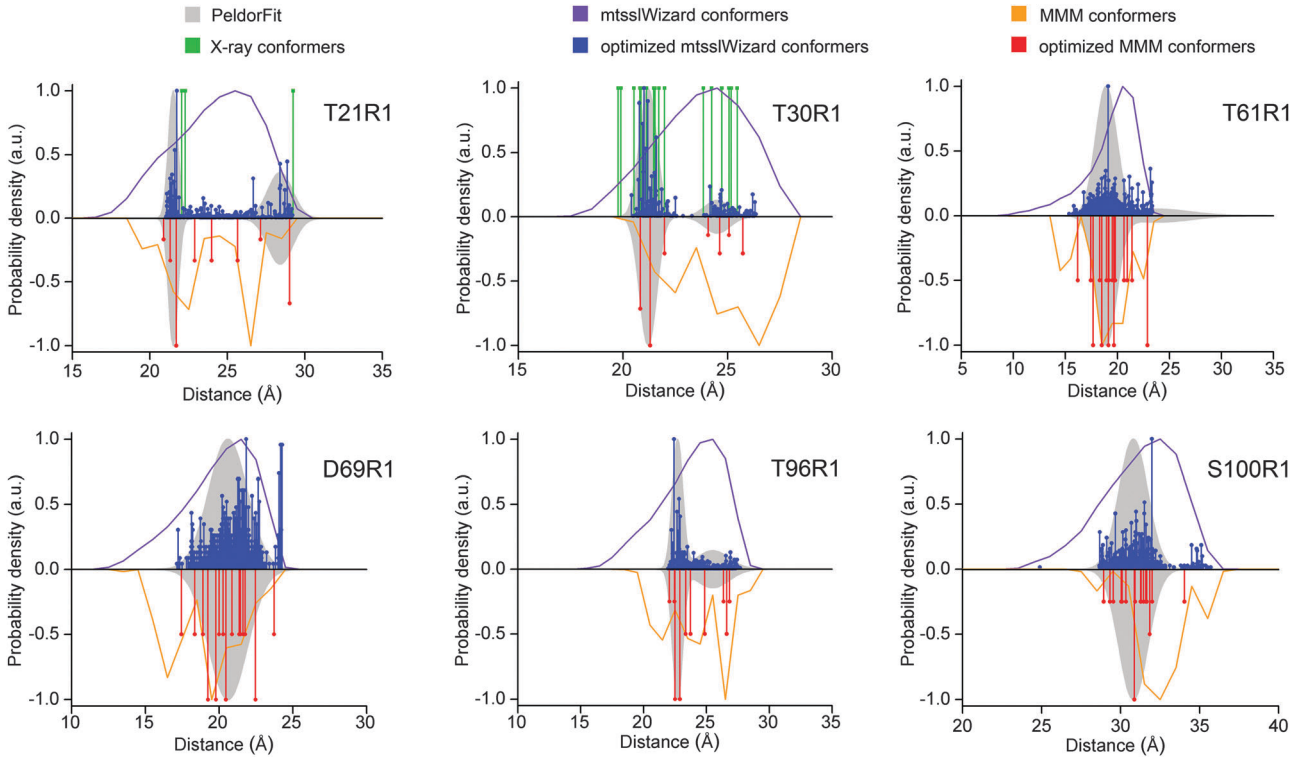


Fig. 4 Cu^{2+}–NO distance distributions for the six azurin mutants. The experimental Cu^{2+}–NO distance distributions for the six azurin mutants are shown as grey shades. Top-half in each plot: The Cu^{2+}–NO distance distribution for the initial (violet) and optimised (blue) mtsslWizard-generated R1 conformers. The green vertical lines represent Cu^{2+}–NO distances that were determined from crystal structures. Bottom-half of each plot: The Cu^{2+}–NO distance distribution for the initial (orange) and optimised (red) MMM-generated R1 conformers.

PELDOR time traces. To simulate the different levels of occupancy of conformers, each trial conformer was allowed to be picked multiple times. In the following, the focus is laid on mutants T21R1 and T30R1, because crystal structures for these mutants were determined (for other mutants see Fig. S3, ESI†). The optimised R1 ensembles for these mutants are depicted in Fig. 5. They display very similar positions of their NO groups independent of the program used to generate the input ensemble,

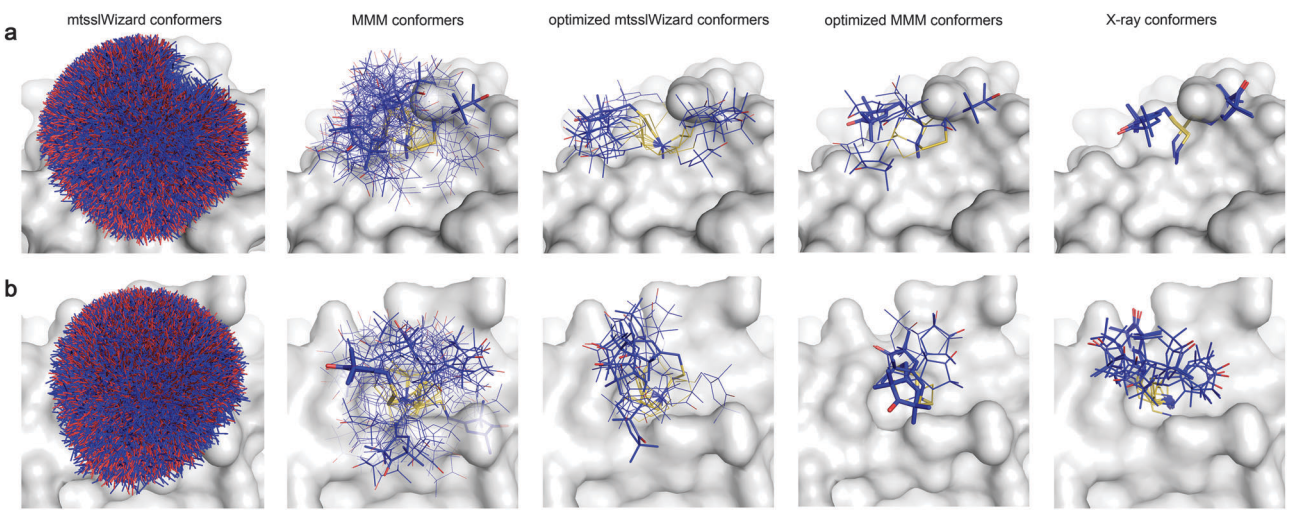


Fig. 5 Models of the R1 side chain for the azurin mutants (a) T21R1 and (b) T30R1. From left to right: (1) mtsslWizard-generated R1 conformers, (2) MMM-generated R1 conformers, (3) optimised mtsslWizard-generated R1 conformers, (4) optimised MMM-generated R1 conformers, (5) X-ray R1 conformers. The molecular surface of azurin is shown in grey. The R1 conformers are shown as blue stick models. The thickness of the sticks is proportional to the abundance of the conformers. The spin-carrying oxygen atom is coloured red.

MMM or mtsslWizard. Fig. 3a and b reveal that the PELDOR time traces that were simulated for both sets of optimised conformers are almost indistinguishable from the experimental time traces. Moreover, the Cu^{2+}–NO distance distributions that were calculated for the optimised R1 sub-ensembles have very similar shapes as compared to the experimental distributions: both are bimodal and match in width, position and relative intensity (Fig. 4).

Note that the PELDOR data sets restrain only the position and orientation of the NO group of the R1 side chain, but not its overall geometry. Therefore, several combinations of the five dihedral angles of the R1 tether may result in approximately the same position and orientation of its NO group. This was observed when 100 independent runs of the optimisation algorithm were performed for each mtsslWizard-based R1 ensemble. All solutions contained R1 conformers with different dihedral angles χ_1–χ_5 but very similar positions and orientations of the NO group. Consequently, the Cu^{2+}–nitroxide distance distributions, which were averaged over the 100 optimisation runs, look identical to the distance distributions calculated after a single run (Fig. S3, ESI†). In contrast, the optimisation of the MMM-based R1 ensembles even provided the same conformers, with the same dihedral angles χ_1–χ_5, which is most probably due to a much smaller number of trial conformers for MMM as compared to mtsslWizard (200 and 10 000). This proves that the optimisation procedure provides a stable solution at least with respect to the position and orientation of the NO group for the given set of PELDOR constraints. The selected R1 conformers may be therefore considered as experimental models of the spin label in frozen solution.

X-ray crystal structure of azurin T30R1

The R1 conformations were also determined from the earlier determined crystal structure of the azurin mutant T21R1[24] and the new crystal structure of the mutant T30R1. T30R1 was found to crystallize in two different crystal forms (I and II). The crystal structures were solved by molecular replacement using a model of native azurin as the search model (PDB 1E67). Both crystal forms were indexed with the space group *C*2 but have different crystal packings with 4 (crystal form I) or 16 (crystal form II) monomers in the asymmetric unit. While the overall protein structures of all monomers are virtually identical, different conformations of the R1 side chain were observed in the electron density (Fig. S4, ESI†). The quality of the electron density for different R1 conformers ranged from excellent (crystal form – chain: I-A, I-B, II-A, II-B, II-C, II-E, II-F, II-K, II-L, II-P) over average (I-D, II-G, II-H, II-I, II-M, II-N, II-O) to poor (I-C, II-D, II-J) (Fig. S5a and Table S3, ESI†). Nevertheless, in each case the overall conformation of the spin label and thus the position of the NO group could be determined.

The obtained crystal structures reveal that the R1 side chain is involved in multiple, mostly hydrophobic interactions with the "parent" azurin molecule and also in a varying number of crystal contacts (Fig. S5b, ESI†). In order to check how these interactions influence the geometry of the R1 side chain, the five dihedral angles χ_1–χ_5 (see definitions in Fig. 1) were determined for each R1 conformer (Fig. 6 and Table S3, ESI†). Obviously, the precision of these angles depends on the quality of the electron density. Here, all five dihedral angles could be determined for the four R1 conformers of T21R1 and for 10 out of 20 R1 conformers of T30R1. The obtained dihedral angles were then compared with the canonical angles (χ_1, χ_2, χ_3,: -60°, 60°, 180°; χ_3, χ_5: -90°, 90°) from rotamer libraries (Fig. 6 and Table S3, ESI†). In analogy to Hubbell and co-workers,[17] a $\pm 30^\circ$ confidence interval was assigned to each canonical angle. In the case of T21R1, two out of four R1 conformers have dihedral angles, which fall within the confidence intervals

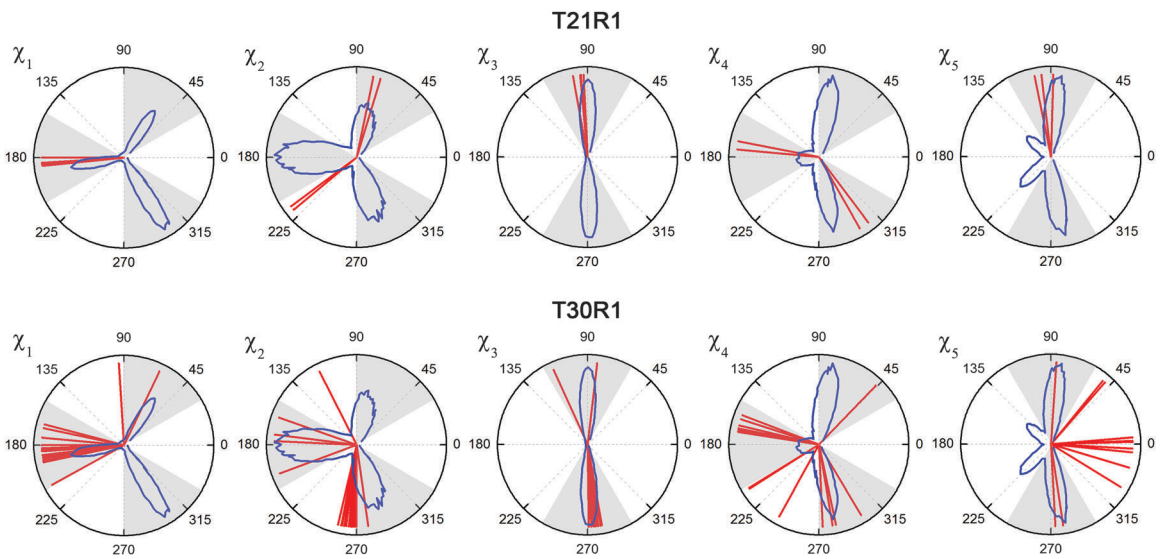


Fig. 6 The dihedral angles χ_1–χ_5 of the R1 side chain: X-ray structures of in T21R1 and T30R1 (red sticks), canonical angles with the $\pm 30^\circ$ confidence interval (grey shades), MMM rotamer library (blue lines).

of the canonical angles, whereas the other two display a 40° deviation of their χ_2 value from the closest canonical angle. Also in all R1 conformers of T30R1 at least one dihedral angle deviates from the corresponding canonical angle by more than 30°. The obtained deviations of individual dihedral angles from the values corresponding to a free spin label in the rotamer libraries reveal that the interactions of the spin labels with their protein environments can affect the conformation of the R1 side chain.

Comparison of the R1 models for the frozen solution with the R1 crystal structures

In Fig. 5, the crystallographic R1 conformations of T21R1 and T30R1 are compared with the corresponding PELDOR-derived models of R1 in frozen solution. Interestingly, in spite of the R1 side chain's' involvement in crystal contacts, the comparison reveals significant similarities between the crystal structures and the models with respect to the location and orientation of the NO group. In the case of mutant T21R1 two distinct R1 sub-ensembles can be identified (Fig. 5a), the corresponding Cu^{2+}–NO distances of ~22 and 28 Å do match the PELDOR-derived distances (Fig. 4), and the relative integral weights of the two sub-ensembles are also very similar for both, the frozen solution and the crystal. In the case of the T30R1 mutant, the correlation between the crystal structures and the models is less pronounced as compared to T21R1, but still the overall positions of the R1 ensembles match (Fig. 5b). Here, it is important to note that the conformations of the R1 side chain are different for the two crystal forms (I and II) of T30R1 (Fig. 7). On the one hand, the positions of the NO group in the crystal form I (Fig. 7a) are very similar to the ones in the optimised mtsslWizard and MMM models of R1 (Fig. 5b). As a consequence, these conformers reproduce well the main peak in the PELDOR-derived distance distribution, appearing at 21 Å (Fig. 6c). On the other hand, none of the crystal form I conformers resemble the low-intensity peak at 25 Å. In contrast, the R1 conformations found in crystal form II (Fig. 7b) reflect the bimodality of the PELDOR-derived distance distribution (Fig. 7c). Only the relative intensities of the two maxima are different, 5 : 3 in the crystal and 5 : 1 in frozen solution. Adding the R1 ensembles from both crystal forms increases the ratio to 7 : 3, leveraging the difference to some extent.

The similarity of the orientations of the NO group in the crystals and frozen solutions of T21R1 and T30R1 is also seen when simulating the PELDOR time traces using the crystallographic R1 conformers (Fig. 3c). The simulated time traces do not provide as good fits to the experimental time traces as the optimized *in silico* models, but the agreement is still very good. This is reflected in very similar modulation depths and modulation frequencies. In contrast, the original MMM- and mtsslWizard-based R1 ensembles do not yield time traces that fit as well as the time traces from the crystallographic R1 ensembles.

Despite the similarity in the positions of the NO group, the optimised mtsslWizard- and MMM-based conformers deviate from each other and from the X-ray R1 conformers by their five dihedral angles (Fig. 8). This may have several reasons. First, even though the optimisation of the *in silico* R1 ensembles provides stable solutions, the limited orientation selectivity of the X-band PELDOR data sets leaves some uncertainty with respect to the values of the individual dihedral angles. Second, several combinations of dihedral angles may yield approximately the same position and orientation of the NO group. Third, some of the crystallographic R1 conformers do not appear in the rotamer library used by MMM (Fig. 6), which is probably due to the complex interactions of the spin label with the protein surrounding.

In summary, the NO group of the R1 side chain of azurin mutants T21R1 and T30R1 is shown to have very similar positions and orientations in frozen solution and in protein crystals. A possible reason for this result maybe that azurin crystals contain 50% solvent as typically found for protein crystals,[45] and this solvent also surrounds a large proportion of the label, making both environments, crystal and frozen solution, not so different. However, it is important to note that R1 conformations that have low abundances in frozen solution can be absent or, opposite, abundant in the crystals, as it is found for T30R1. The latter result shows that in some cases only a fraction of the R1 conformations that occur in solution is selected during the crystal formation. Which conformations are selected will depend on the crystallization conditions and crystal packing. In essence, this means that a single crystal structure of an R1 side chain at a particular labelling site will likely give a correct snap shot of R1 in solution, whereas many structures in different crystal packings (or large asymmetric units) are needed to get a complete picture of the R1 side chain in frozen solution. Moreover, the conformational ensemble of R1 in frozen glassy solution will depend on the freezing point of the solvent/buffer.

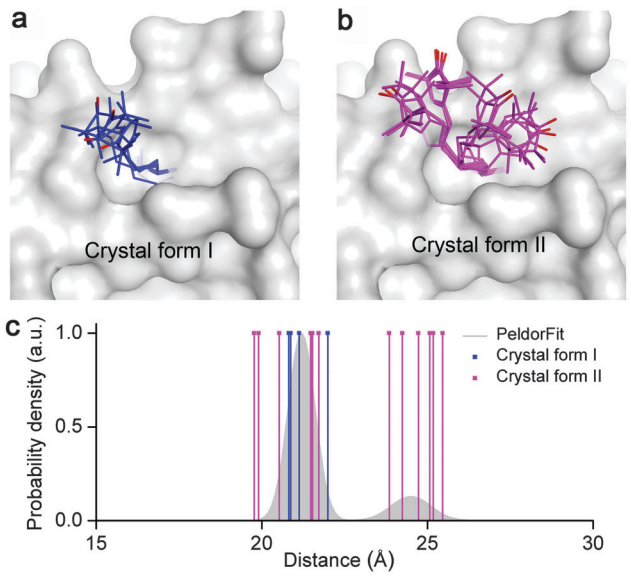


Fig. 7 The conformations of the R1 side chain in the T30R1 crystal structures of (a) form I and (b) form II. (c) The Cu^{2+}–NO distance distributions calculated for two crystal structures of T30R1. The experimental Cu^{2+}–NO distance distribution of T30R1 is shown as grey shade.

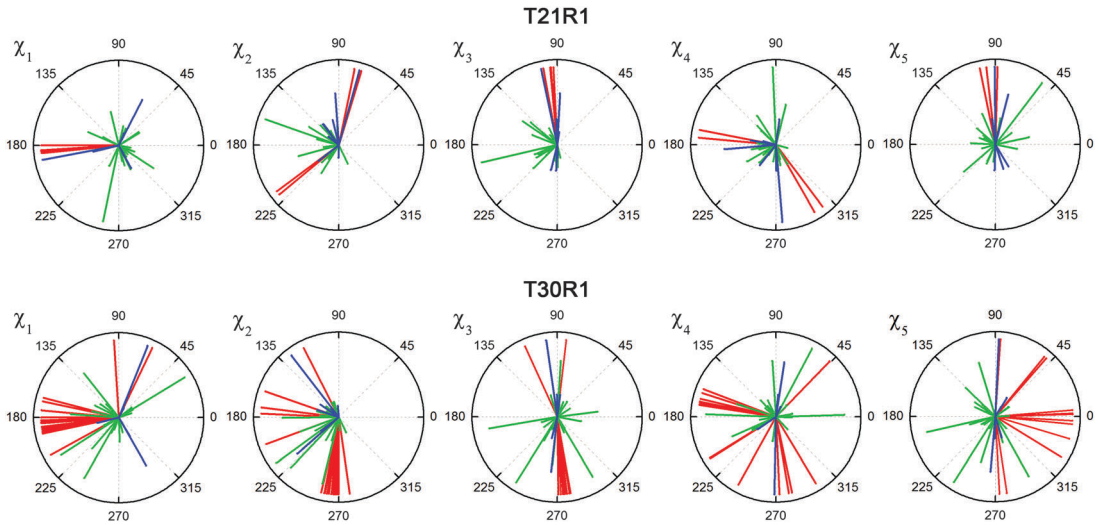


Fig. 8 The dihedral angles χ_1–χ_5 of the R1 side chain in T21R1 and T30R1: X-ray R1 structures (red), optimised MMM conformers (blue), optimised mtsslWizard conformers (green).

Conclusions

Structural models of the R1 side chain in frozen solution were derived from orientation selective PELDOR data of azurin. Both, the accessible volume approach (mtsslWizard) and the rotamer approach (MMM) were used to sample the conformational space of R1 and were both found to predict much broader distance distributions than what was observed experimentally. However, after optimization based on PELDOR constraints, both *in silico* methods yielded very similar conformational ensembles. In the case of the T21R1 and T30R1 mutants, these models were compared to the corresponding crystal structures, revealing significant similarities in the NO positions and orientations. Although the crystallization conditions may favour certain conformations present in frozen glassy solution as found for T30R1.

Current *in silico* spin labelling algorithms are indispensable tools for the interpretation of EPR-based distance data and sometimes they show very good agreement with the experiments, especially at solvent exposed sites, when the label can freely sample its accessible volume. However, the labelling site cannot always be chosen freely and often, more occluded sites have to be used. Thus, the aim should be to improve the prediction accuracy independent of the labelling site. Also, as long as the precision of high-quality PELDOR distance measurements is higher than the prediction accuracy of *in silico* spin labelling methods, it is important to improve these algorithms in order to reduce the number of distance constraints needed for coarse-grained modelling, docking, trilateration or for monitoring small structural changes. The fact that algorithms of different sophistication produce very similar results proves that it is not sufficient to include only selected contributors to the energetic state of the spin label in its environment. We think that in the long run, only more holistic approaches such as full-atom MD or QM/MM, which include conformational changes of the protein at the label site, label-protein interactions, label-solvent and label-solute interactions as well as temperature dependencies, can succeed to level the discrepancy between experiment and prediction. The presented models and crystal structures of the azurin mutants provide important information on the way to improve such methods.

Acknowledgements

This work was funded by the German Research Foundation (DFG) with a grant for project A6 within the collaborative research centre SFB813 ''Chemistry at Spin Centres''.

References

1 D. Milov, K. M. Salikhov and M. D. Shirov, *Fiz. Tverd. Tela*, 1981, **23**, 975; R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich and H. W. Spiess, *Angew. Chem., Int. Ed.*, 1998, **37**, 2833.
2 G. Jeschke, *Annu. Rev. Phys. Chem.*, 2012, **63**, 419; O. Schiemann and T. F. Prisner, *Q. Rev. Biophys.*, 2007, **40**, 1.
3 Z. Yang, Y. Liu, P. Borbat, J. L. Zweier, J. H. Freed and W. L. Hubbell, *J. Am. Chem. Soc.*, 2012, **134**, 9950; G. Yu. Shevelev, O. A. Krumkacheva, A. A. Lomzov, A. A. Kuzhelev, O. Yu. Rogozhnikova, D. V. Trukhin, T. I. Troitskaya, V. M. Tormyshev, M. V. Fedin, D. V. Pyshnyi and E. G. Bagryanskaya, *J. Am. Chem. Soc.*, 2014, **136**, 9874.
4 I. Krstić, R. Hänsel, O. Romainczyk, J. W. Engels, V. Dötsch and T. F. Prisner, *Angew. Chem., Int. Ed.*, 2011, **50**, 5070; M. Qi, A. Groß, G. Jeschke, A. Godt and M. Drescher, *J. Am. Chem. Soc.*, 2014, **136**, 15366.
5 D. Goldfarb, in *Structure and Bonding*, ed. C. R. Timmel and J. R. Harmer, Springer-Verlag, Berlin/Heidelberg, 2013, vol. 152, pp. 163–204; E. Narr, A. Godt and G. Jeschke, *Angew. Chem., Int. Ed.*, 2002, **41**, 3907.
6 M. Bennati, A. Weber, J. Antonic, D. L. Perlstein, J. Robblee and J. Stubbe, *J. Am. Chem. Soc.*, 2003, **125**, 14988; V. P. Denysenkov,

T. F. Prisner, J. Stubbe and M. Bennati, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13386.
7 J. P. Klare and H.-J. Steinhoff, in *Structure and Bonding*, ed. C. R. Timmel and J. R. Harmer, Springer-Verlag, Berlin/Heidelberg, 2013, vol. 152, pp. 205–248; A. D. Milov, A. G. Maryasov, Y. D. Tsvetkov and J. Raap, *Chem. Phys. Lett.*, 1999, **303**, 135.
8 H. El Mkami and D. G. Norman, *Methods Enzymol.*, 2015, **564**, 125.
9 P. P. Borbat, K. Surendhran, M. Bortolus, P. Zou, J. H. Freed and H. S. Mchaourab, *PLoS Biol.*, 2007, **5**, e271.
10 C. Pliotas, R. Ward, E. Branigan, A. Rasmussen, G. Hagelueken, H. Huang, S. S. Black, I. R. Booth, O. Schiemann and J. H. Naismith, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, E2675.
11 D. Abdullin, N. Florin, G. Hagelueken and O. Schiemann, *Angew. Chem., Int. Ed.*, 2015, **54**, 1827.
12 N. Alexander, M. Bortolus, A. Al-Mestarihi, H. Mchaourab and J. Meiler, *Structure*, 2008, **16**, 181; O. Duss, E. Michel, M. Yulikov, M. Schubert, G. Jeschke and F. H.-T. Allain, *Nature*, 2014, **509**, 588; R. Sircar, P. P. Borbat, M. J. Lynch, J. Bhatnagar, M. S. Beyersdorf, C. J. Halkides, J. H. Freed and B. R. Crane, *J. Mol. Biol.*, 2015, **427**, 867; C. M. Hammond, T. Owen-Hughes and D. G. Norman, *Methods*, 2014, **70**, 139.
13 L. J. Berliner, J. Grunwald, H. O. Hankovszky and K. Hideg, *Anal. Biochem.*, 1982, **119**, 450.
14 C. Abé, D. Klose, F. Dietrich, W. H. Ziegler, Y. Polyhach, G. Jeschke and H.-J. Steinhoff, *J. Magn. Reson.*, 2012, **216**, 53.
15 B. Endeward, J. A. Butterwick, R. MacKinnon and T. F. Prisner, *J. Am. Chem. Soc.*, 2009, **131**, 15246.
16 B. H. Robinson, L. J. Slutsky and F. P. Auteri, *J. Chem. Phys.*, 1992, **96**, 2609; H. J. Steinhoff and W. L. Hubbell, *Biophys. J.*, 1996, **71**, 2201; D. E. Budil, K. L. Sale, K. A. Khairy and P. G. Fajer, *J. Phys. Chem. A*, 2006, **110**, 3703; F. Tombolato, A. Ferrarini and J. H. Freed, *J. Phys. Chem. B*, 2006, **110**, 26248; S. C. DeSensi, D. P. Rangel, A. H. Beth, T. P. Lybrand and E. J. Hustedt, *Biophys. J.*, 2008, **94**, 3798; V. S. Oganesyan, *Phys. Chem. Chem. Phys.*, 2011, **13**, 4724.
17 Z. Guo, D. Cascio, K. Hideg, T. Kálái and W. L. Hubbell, *Protein Sci.*, 2007, **16**, 1069; Z. Guo, D. Cascio, K. Hideg K, T. Kálái and W. L. Hubbell, *Protein Sci.*, 2008, **17**, 228; M. R. Fleissner, D. Cascio and W. L. Hubbell, *Protein Sci.*, 2009, **18**, 893.
18 G. Hagelueken, W. J. Ingledew, H. Huang, B. Petrovic-Stojanovska, C. Whitfield, H. ElMkami, O. Schiemann and J. H. Naismith, *Angew. Chem., Int. Ed.*, 2009, **48**, 2904.
19 D. M. Freed, P. S. Horanyi, M. C. Wiener and D. S. Cafiso, *Biophys. J.*, 2010, **99**, 1604.
20 J. E. D. Lillington, J. E. Lovett, S. Johnson, P. Roversi, C. R. Timmel and S. M. Lea, *J. Mol. Biol.*, 2011, **405**, 427.
21 T. Gruene, M. K. Cho, I. Karyagina, H. Y. Kim, C. Grosse, K. Giller, M. Zweckstetter and S. Becker, *J. Biomol. NMR*, 2011, **49**, 111.
22 T. F. Cunningham, M. S. McGoff, I. Sengupta, C. P. Jaroniec, W. S. Horne and S. Saxena, *Biochemistry*, 2012, **51**, 6350.
23 S. Stoll, Y. T. Lee, M. Zhang, R. F. Wilson, R. D. Britt and D. B. Goodin, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 12888.
24 N. Florin, O. Schiemann and G. Hagelueken, *BMC Struct. Biol.*, 2014, **14**, 16.
25 M. Plato, H.-J. Steinhoff, C. Wegener, J. T. Törring, A. Savitsky and K. Möbius, *Mol. Phys.*, 2002, **100**, 3711; L. Urban and H.-J. Steinhoff, *Mol. Phys.*, 2013, **111**, 18; P. Gast, R. T. L. Herbonnet, J. Klare, A. Nalepa, C. Rickert, D. Stellinga, L. Urban, K. Möbius, A. Savitsky, H.-J. Steinhoff and E. J. J. Groenen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 15910.
26 O. Schiemann, N. Piton, Y. Mu, G. Stock, J. W. Engels and T. F. Prisner, *J. Am. Chem. Soc.*, 2004, **126**, 5722.
27 K. Sale, L. Song, Y.-S. Liu, E. Perozo and P. Fajer, *J. Am. Chem. Soc.*, 2005, **127**, 9334.
28 C. Beier and H. J. Steinhoff, *Biophys. J.*, 2006, **91**, 2647; M. I. Fajer, H. Li, W. Yang and P. G. Fajer, *J. Am. Chem. Soc.*, 2007, **129**, 13840; D. Sezer, J. H. Freed and B. Roux, *J. Phys. Chem. B*, 2008, **112**, 5755; D. Sezer, J. H. Freed and B. Roux, *J. Am. Chem. Soc.*, 2009, **131**, 2597; J. L. Sarver, J. E. Townsend, G. Rajapakse, L. Jen-Jacobson and S. Saxena, *J. Phys. Chem. B*, 2012, **116**, 4024; V. Barone, P. Cimino and E. Stendardo, *J. Chem. Theory Comput.*, 2008, **4**, 751; E. Stendardo, A. Pedone, P. Cimino, M. C. Menziani, O. Crescenzi and V. Barone, *Phys. Chem. Chem. Phys.*, 2010, **12**, 11697.
29 G. Hagelueken, R. Ward, J. Naismith and O. Schiemann, *Appl. Magn. Reson.*, 2012, **42**, 377; G. Hagelueken, D. Abdullin, R. Ward and O. Schiemann, *Mol. Phys.*, 2013, **111**, 2757; G. Hagelueken, D. Abdullin and O. Schiemann, *Methods Enzymol.*, 2015, **563**, 595.
30 Y. Polyhach, E. Bordignon and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2356.
31 D. Klose, J. P. Klare, D. Grohmann, C. W. M. Kay, F. Werner and H. J. Steinhoff, *PLoS One*, 2012, **7**, e39492; M. M. Hatmal, Y. Li, B. G. Hegde, P. B. Hegde, C. C. Jao, R. Langen and I. S. Haworth, *Biopolymers*, 2012, **97**, 35; N. S. Alexander, R. A. Stein, H. A. Koteiche, K. W. Kaufmann, H. S. Mchaourab and J. Meiler, *PLoS One*, 2013, **8**, e72851; G. Jeschke, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2013, **72**, 42.
32 S. M. Islam and B. Roux, *J. Phys. Chem. B*, 2015, **119**, 3901.
33 G. Jeschke and Y. Polyhach, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1895.
34 W. A. Coremans, O. G. Poluektov, E. J. J. Groenen, G. W. Canters, H. Nar and A. Messerschmidt, *J. Am. Chem. Soc.*, 1994, **116**, 3097; J. E. Roberts, J. F. Cline, V. Lum, H. B. Gray, H. Freeman, J. Peisach, B. Reinhammar and B. M. Hoffman, *J. Am. Chem. Soc.*, 1984, **106**, 5324; W. E. Antholine, P. M. Hanna and D. R. McMillin, *Biophys. J.*, 1993, **64**, 267.
35 T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. W. Leslie, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 271.
36 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658.
37 H. Nar, R. Huber, A. Messerschmidt, A. C. Filippou, M. Barth, M. Jaquinod and M. Van De Kamp, *Eur. J. Biochem.*, 1992, **205**, 1123.

38 P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 213.
39 V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 12.
40 D. Margraf, P. Cekan, T. F. Prisner, S. Th. Sigurdsson and O. Schiemann, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6708.
41 D. Abdullin, G. Hagelueken, R. I. Hunter, G. M. Smith and O. Schiemann, *Mol. Phys.*, 2015, **113**, 544.
42 P. Bernadó, E. Mylonas, M. V. Petoukhov, M. Blackledge and D. I. Svergun, *J. Am. Chem. Soc.*, 2007, **129**, 5656.
43 D. T. Warshaviak, L. Serbulea, K. N. Houk and W. L. Hubbell, *J. Phys. Chem. B*, 2011, **115**, 397.
44 We thank reviewer 1 for hinting us at this.
45 B. W. Matthews, *J. Mol. Biol.*, 1968, **33**, 491.

Determination of nitroxide spin label conformations via PELDOR and X-ray crystallography

D. Abdullin,[‡a] G. Hagelueken[‡a] and O. Schiemann*[a]

[a] *Institute of Physical and Theoretical Chemistry, Wegelerstr. 12, University of Bonn, 53115 Bonn, Germany*

Supporting Information

Table of Contents

Table S1. Experimental parameters of the PELDOR experiments on the azurin mutants.

Mutant	$\Delta\nu$ (MHz)	ν_{det} (GHz)	ν_{pump} (GHz)	B_0 (T)	$t_{\pi/2}$ (ns)	t_{π} (ns)	t_{pump} (ns)
T21R1	160	9.820069	9.660195	0.3442	16	32	18
	220	9.820435	9.600171	0.3419	16	32	20
	440	9.730215	9.290105	0.3309	16	32	50
	560	9.730316	9.170579	0.3267	16	32	60
T30R1	160	9.720280	9.560145	0.3405	10	20	24
	220	9.720340	9.500095	0.3384	10	20	26
	440	9.720201	9.280175	0.3306	10	20	48
	560	9.720984	9.160650	0.3263	10	20	60
T61R1	160	9.820359	9.660465	0.3441	16	32	18
	220	9.823008	9.602100	0.3419	16	32	20
	440	9.725438	9.285308	0.3308	16	32	46
	560	9.817735	9.257722	0.3297	16	32	48
D69R1	160	9.729975	9.570155	0.3408	16	32	24
	220	9.730461	9.510207	0.3387	16	32	26
	440	9.730301	9.290196	0.3309	16	32	46
	560	9.730095	9.170255	0.3268	16	32	60
T96R1	160	9.720336	9.560065	0.3404	16	32	26
	220	9.720634	9.500550	0.3384	16	32	30
	440	9.720378	9.290358	0.3310	16	32	48
	560	9.720773	9.160600	0.3263	16	32	60
S100R1	160	9.820003	9.660270	0.3442	16	32	18
	220	9.820569	9.600676	0.3419	16	32	22
	440	9.79212	9.371925	0.3338	16	32	44
	560	9.819427	9.258965	0.3298	16	32	46

Table S2. Data collection and refinement statistics for azurin T30R1.

Parameter	Crystal form I	Crystal form II
PDB ID	5I26	5I28
Wavelength (Å)	0.8943	0.8943
Resolution range (Å)	37.36 - 1.89	18.22 - 1.95
Space group	C 1 2 1	C 1 2 1
Unit cell	86.6 68.8 93.4 90 96.8 90	242.5 105.6 100.1 90 114.4 90
Total reflections	364655 (33359)	705691 (68273)
Unique reflections	39652 (3975)	166299 (16352)
Multiplicity	9.1 (8.4)	4.2 (4.2)
Completeness (%)	90.41 (89.72)	99.63 (98.30)
Mean I/sigma(I)	61.22 (5.83)	9.00 (0.69)
Wilson B-factor	19.61	39.70
R-merge	0. 18 (2.78)	0.163 (2.609)
CC1/2	0.99 (0.16)	0.99 (0.21)
R-work	0.193	0.191
R-free	0.243	0.244
Number of non-hydrogen atoms	4342	17050
macromolecules	4031	16180
ligands	4	22
water	307	848
Protein residues	509	2048
RMS (bonds)	0.008	0.008
RMS (angles)	1.15	1.19
Ramachandran favored (%)	98	97
Ramachandran outliers (%)	0	0.1
Clashscore	8.26	6.56
Average B-factor	28.00	48.40
macromolecules	27.80	48.40
ligands	18.80	52.10
solvent	31.10	48.30

* Statistics for the highest-resolution shell are shown in parentheses.

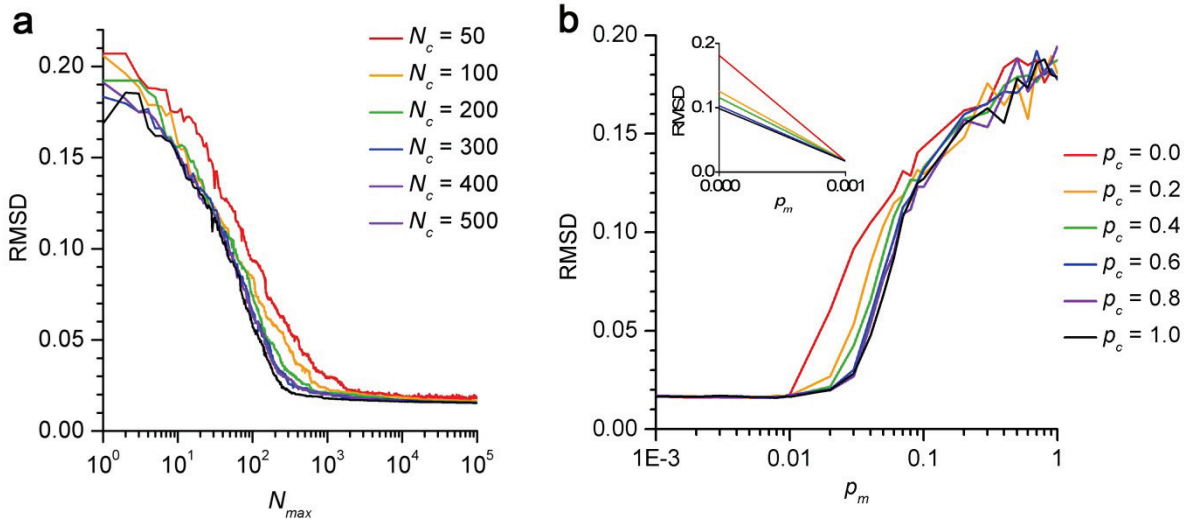


Fig. S1. Tuning the genetic algorithm for the azurin mutant T21R1. RMSD between the experimental and simulated PELDOR signals was calculated as a function of **(a)** N_c and N_{max} setting $N_g = 20$, $p_m = 0.01$, $p_c = 0.5$, **(b)** p_m and p_c setting $N_g = 20$, $N_c = 500$, $N_{max} = 10^4$.

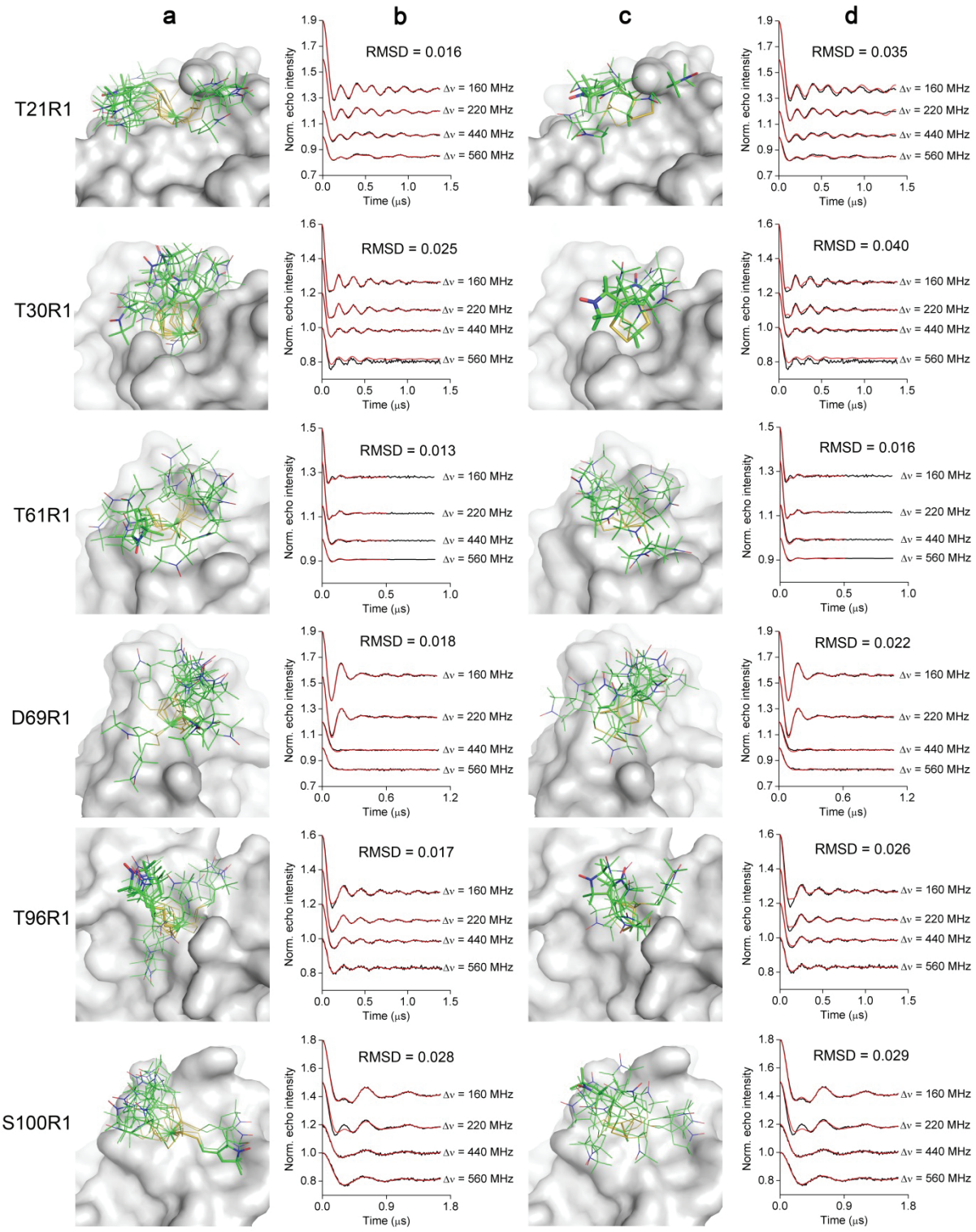


Fig. S2 Results of optimization of the R1 conformers for the six azurin mutants. **(a)** Optimized mtsslWizard-generated R1 conformers and **(b)** simulation (red) of the experimental PELDOR time traces (black) for them. **(c)** Optimized MMM-generated R1 conformers and **(d)** simulation (red) of the experimental PELDOR time traces (black) for them. RMSDs of the simulated PELDOR time traces from the experimental PELDOR time traces are shown on top of the PELDOR time traces.

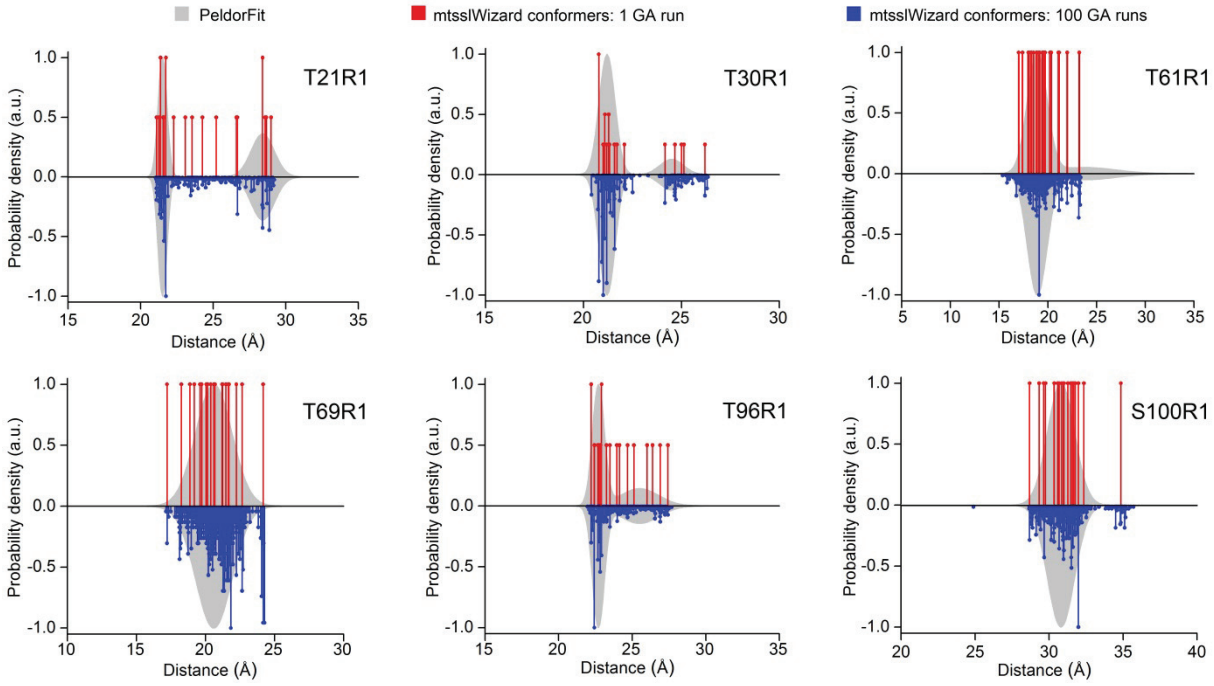


Fig. S3 The Cu^{2+}-nitroxide distance distributions, which are calculated for optimized mtsslWizard-generated R1 conformers after a single run of the genetic algorithm (GA) (red lines) and after 100 consecutive runs (blue lines). The distance distributions from the previous PeldorFit analysis are shown as grey shades.

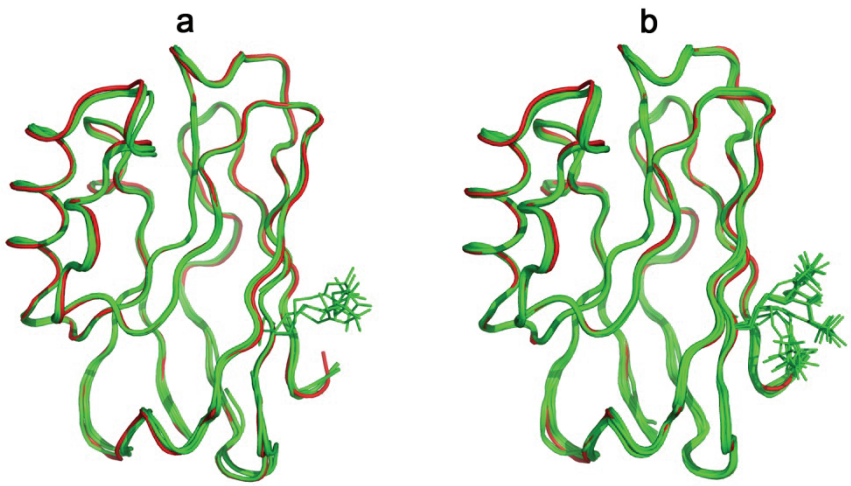
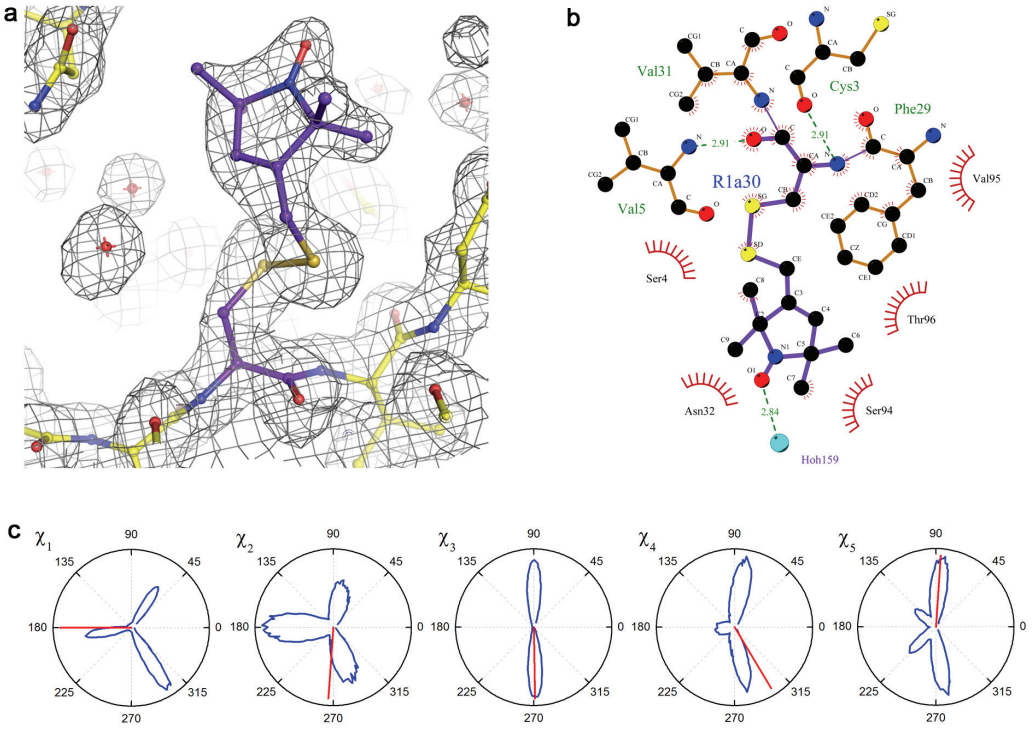


Fig. S4 A superposition of the X-ray crystal structures of the native azurin (PDB 1E67) and azurin mutant T30R1 **(a)** of form A and **(b)** of form B.

T30R1, crystal I, chain A (B-factor = 22.9 Å^2)



T30R1, crystal I, chain B (B-factor = 31.9 Å^2)

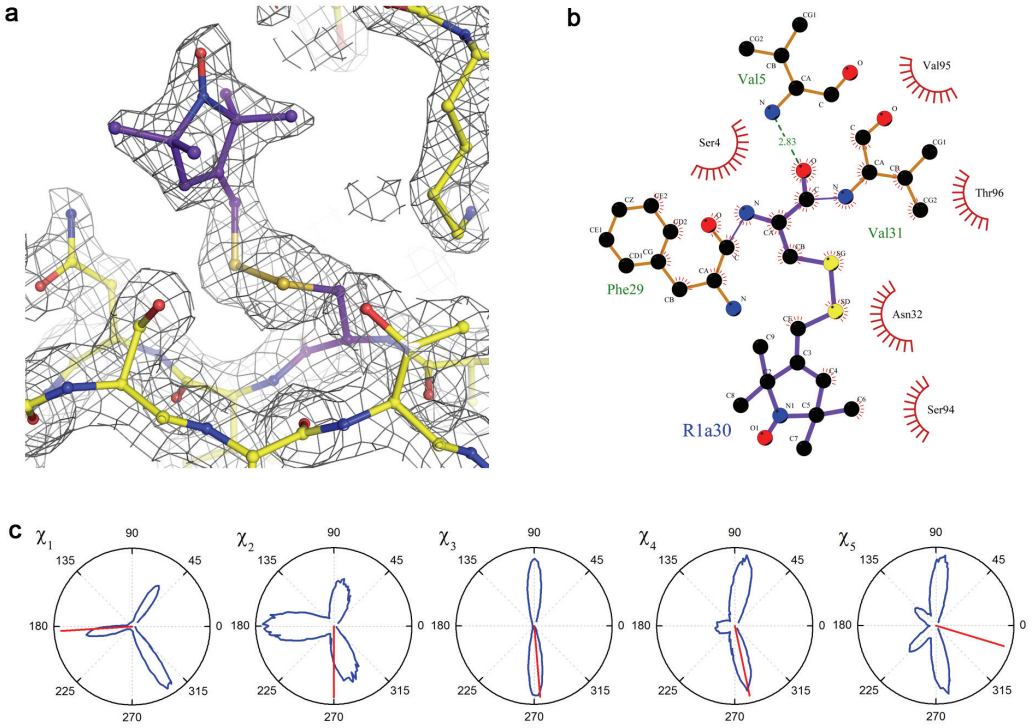
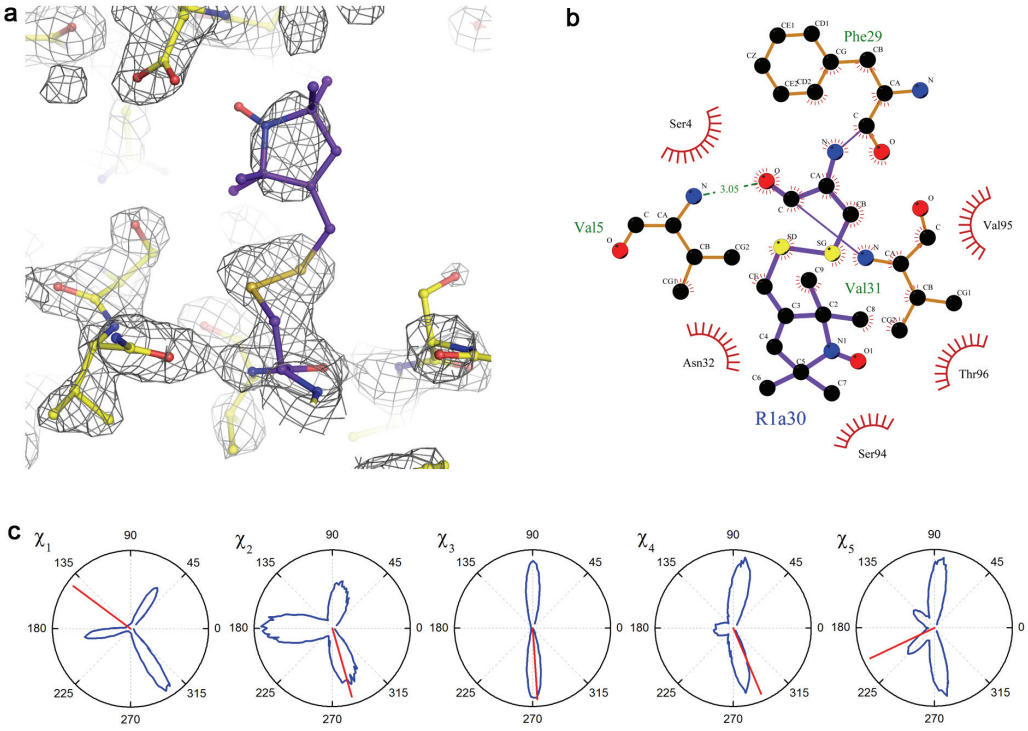


Fig. S5: (caption see below)

T30R1, crystal I, chain C (B-factor = 62.4 Å^2)



T30R1, crystal I, chain D (B-factor = 50.5 Å^2)

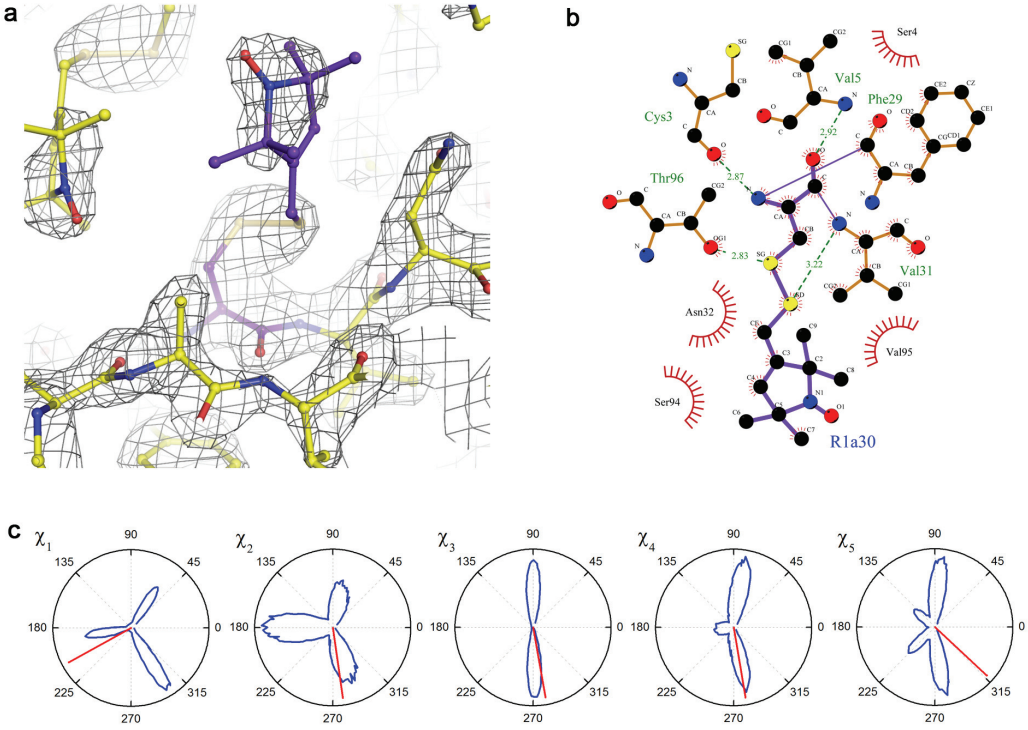
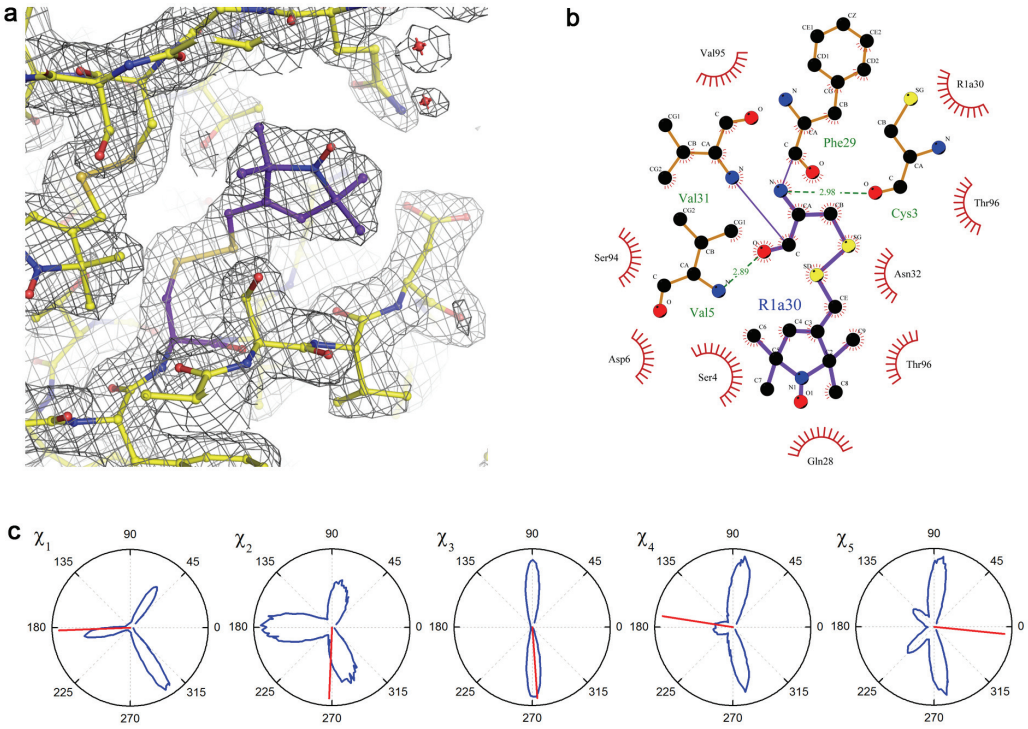


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain A (B-factor = 39.5 Å^2)



T30R1, crystal II, chain B (B-factor = 35.1 Å^2)

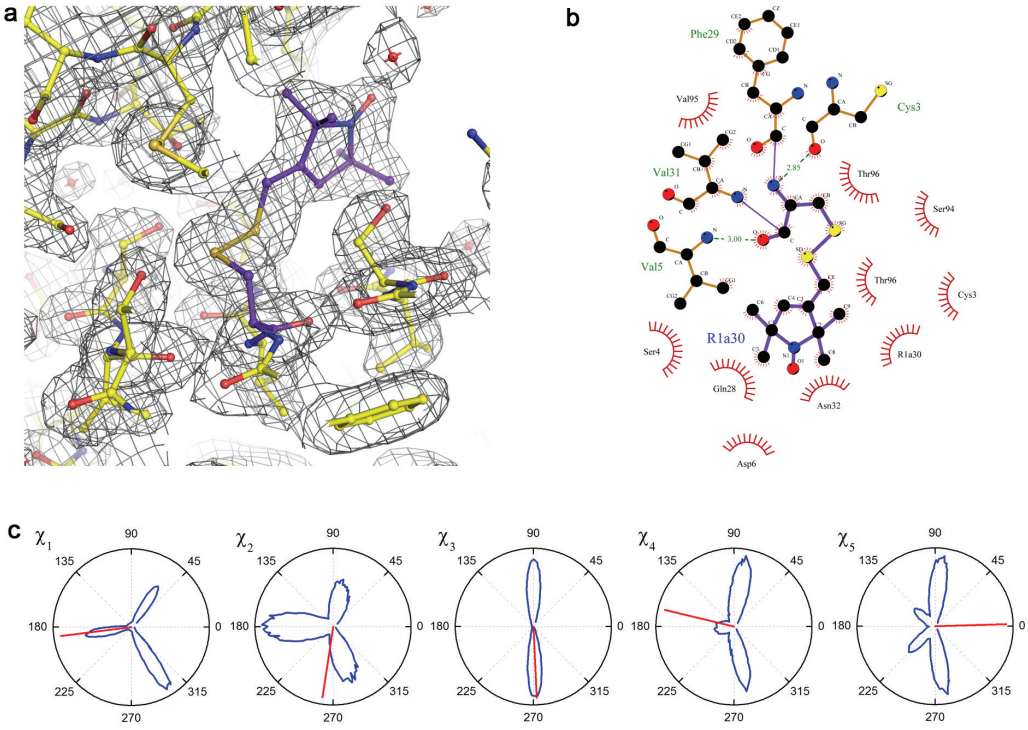
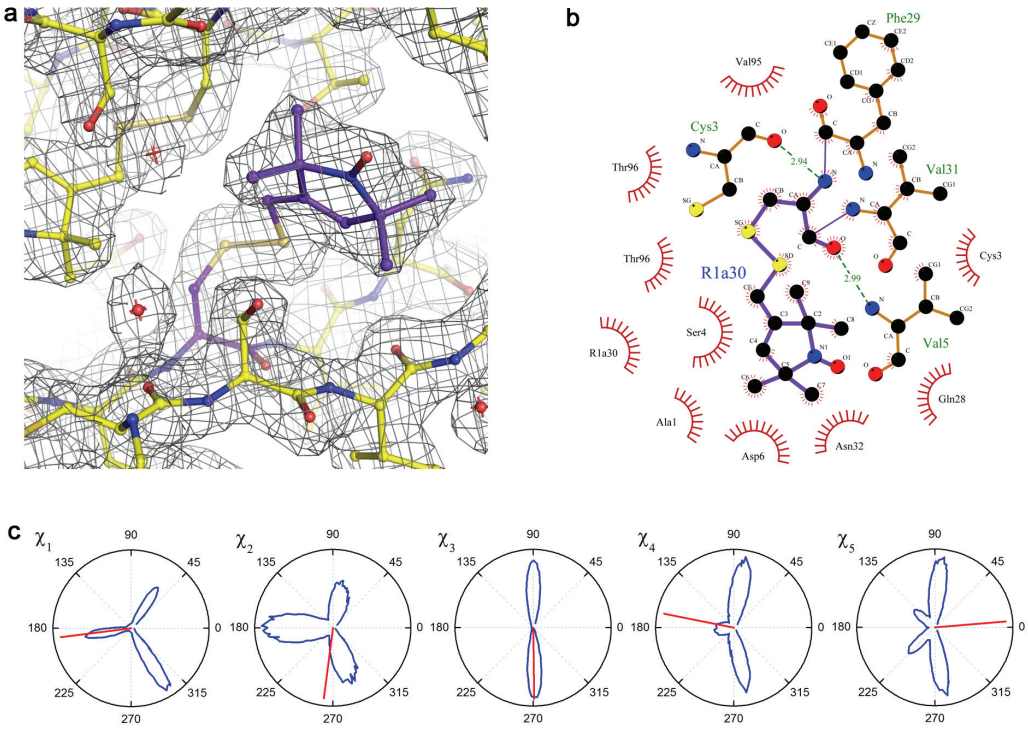


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain C (B-factor = 35.6 Å^2)



T30R1, crystal II, chain D (B-factor = 71.1 Å^2)

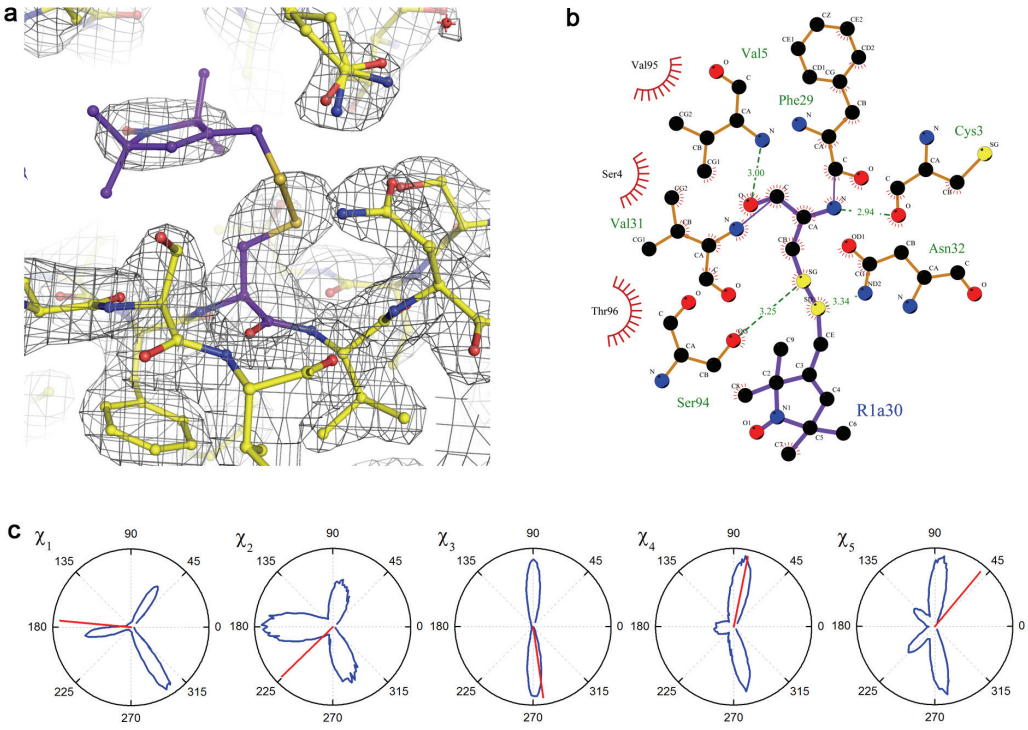
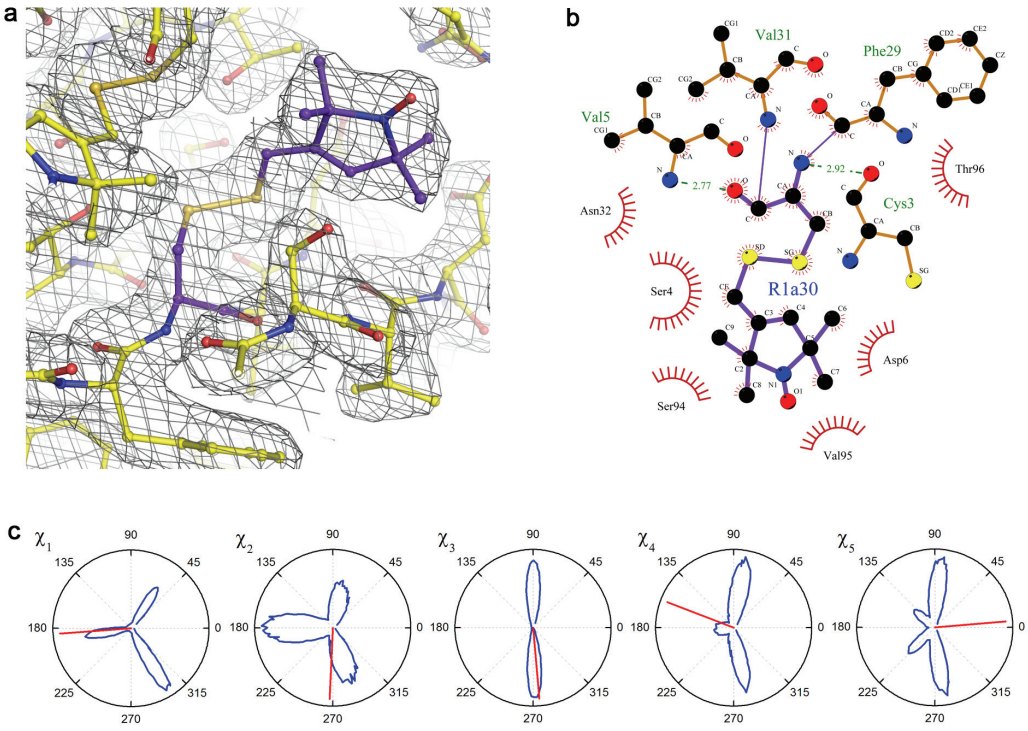


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain E (B-factor = 39.7 Å^2)



T30R1, crystal II, chain F (B-factor = 51.0 Å^2)

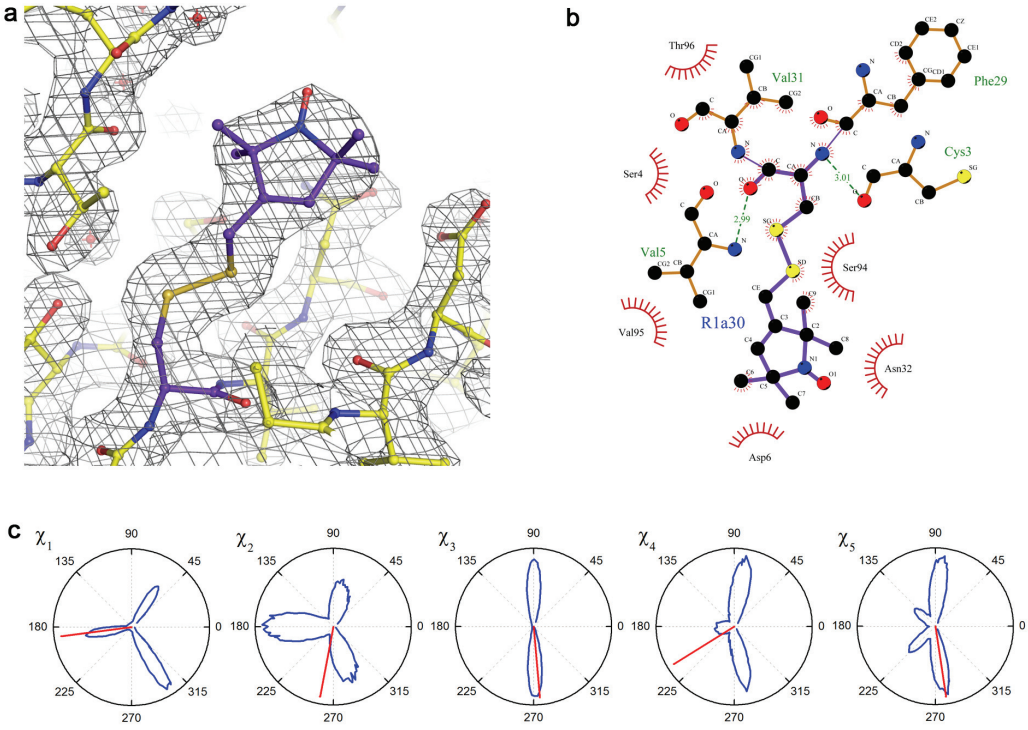
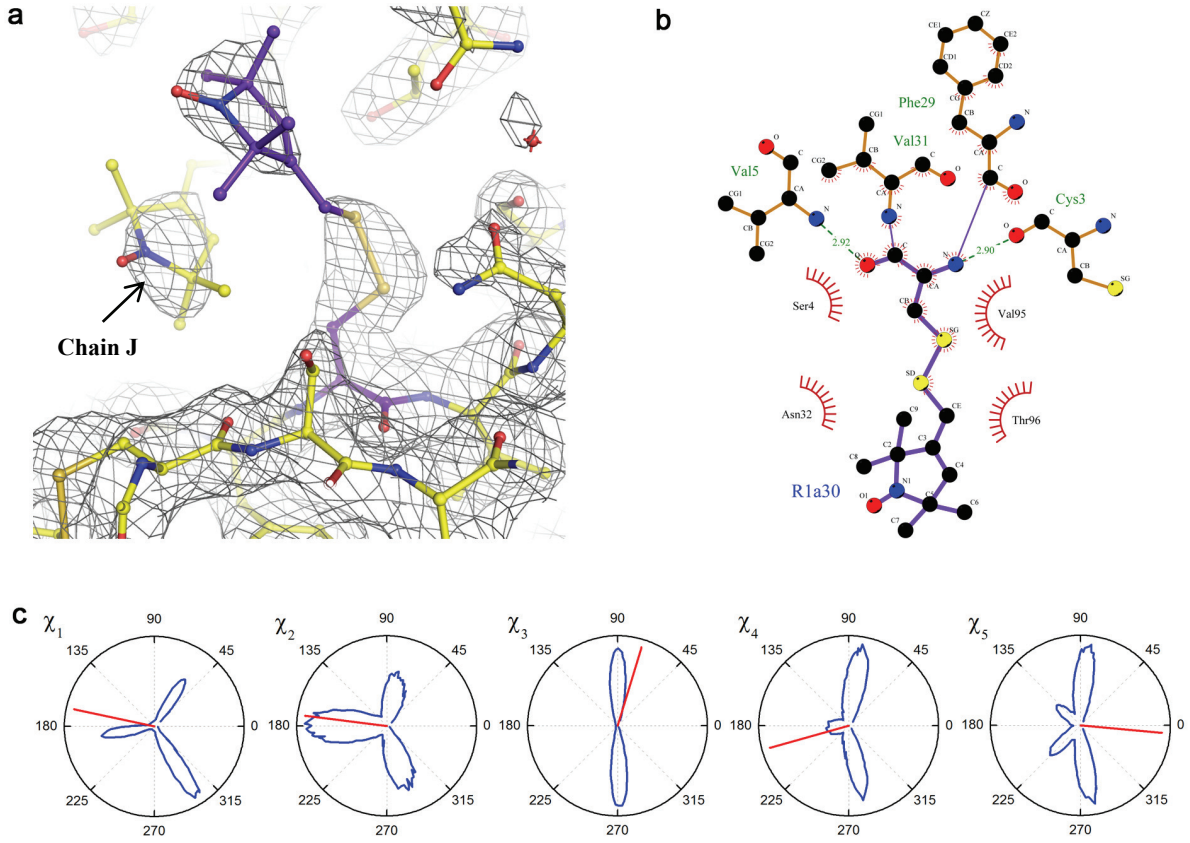


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain G (B-factor = 83.6 Å^2)



T30R1, crystal II, chain H (B-factor = 69.5 Å^2)

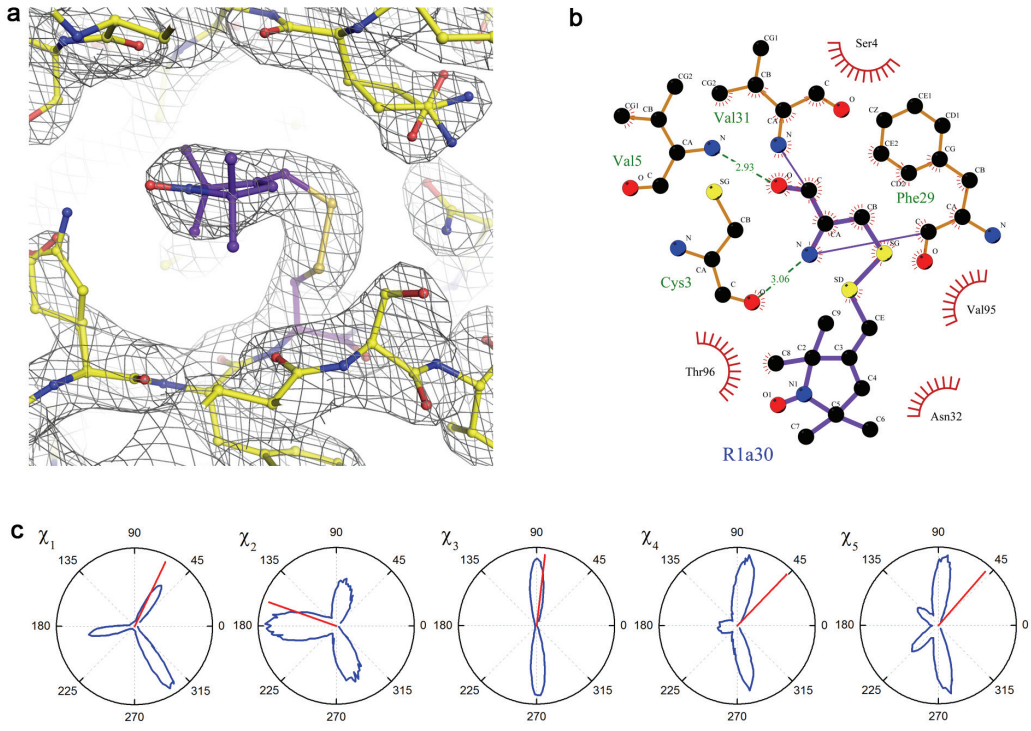
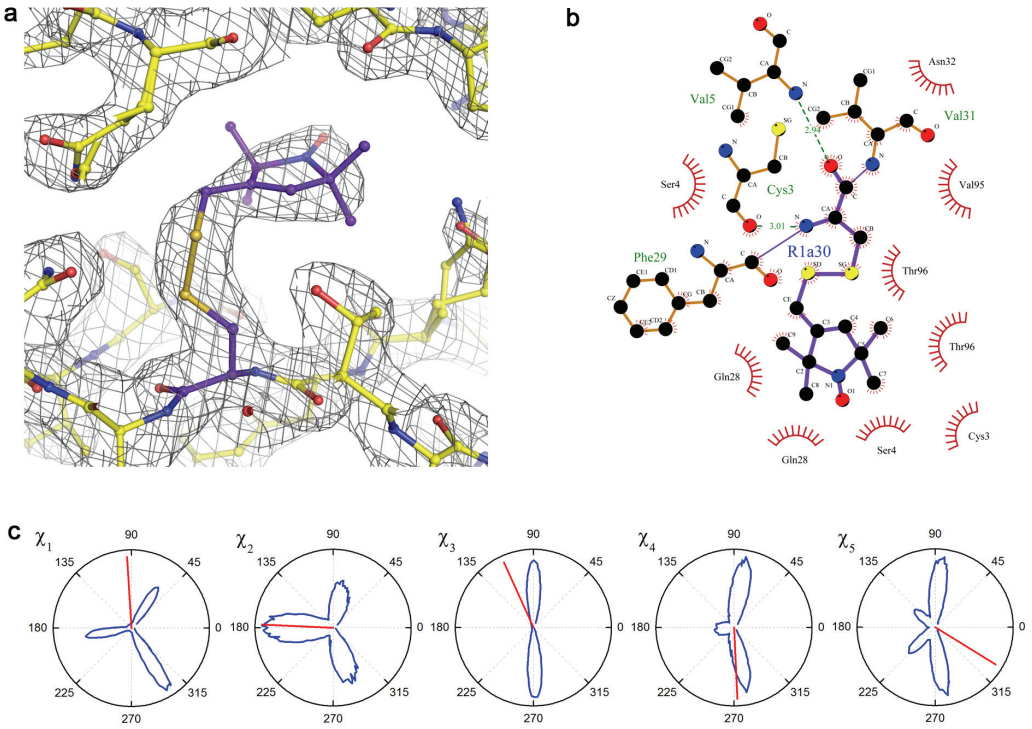


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain I (B-factor = 70.3 Å^2)



T30R1, crystal II, chain J (B-factor = 81.9 Å^2)

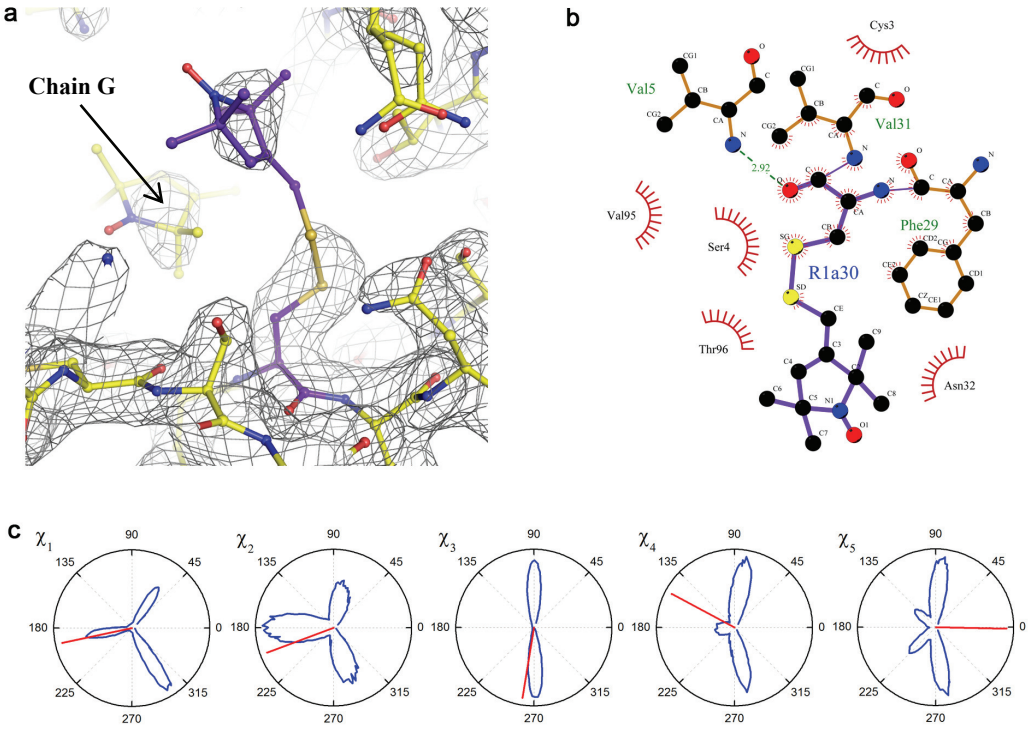
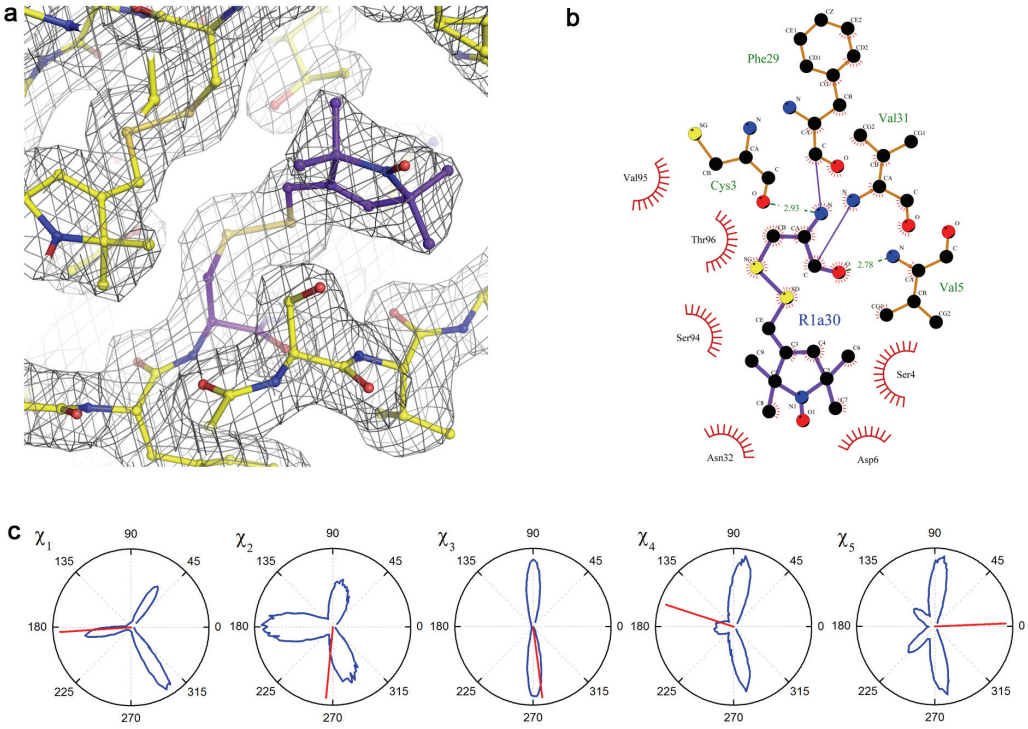


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain K (B-factor = 42.2 Å^2)



T30R1, crystal II, chain L (B-factor = 53.8 Å^2)

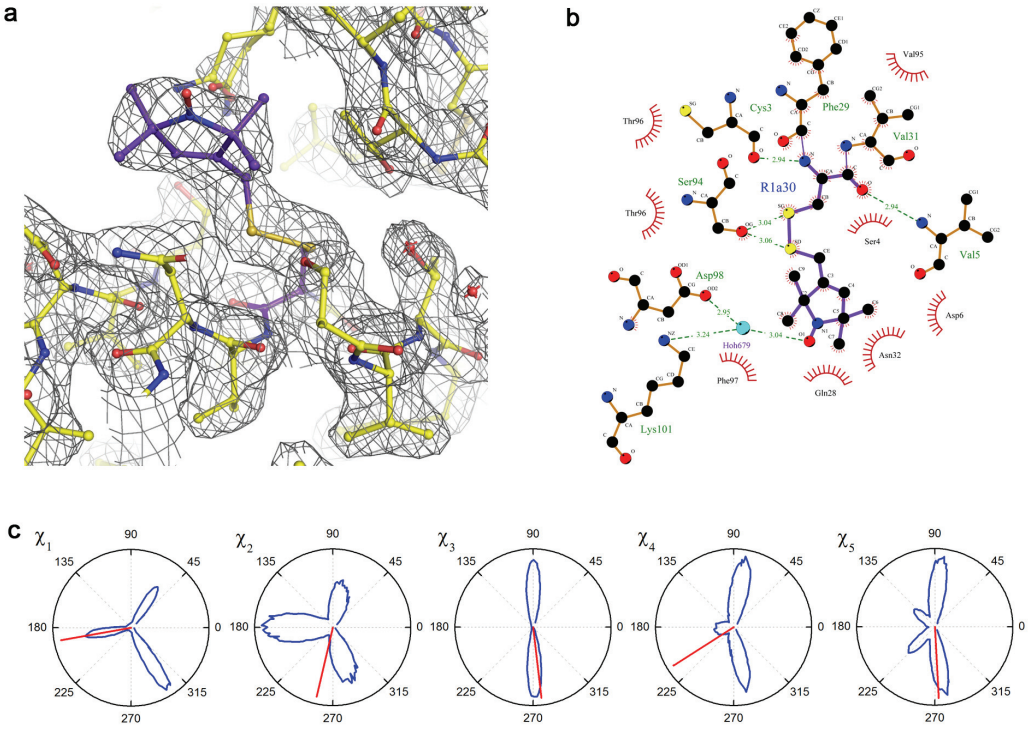
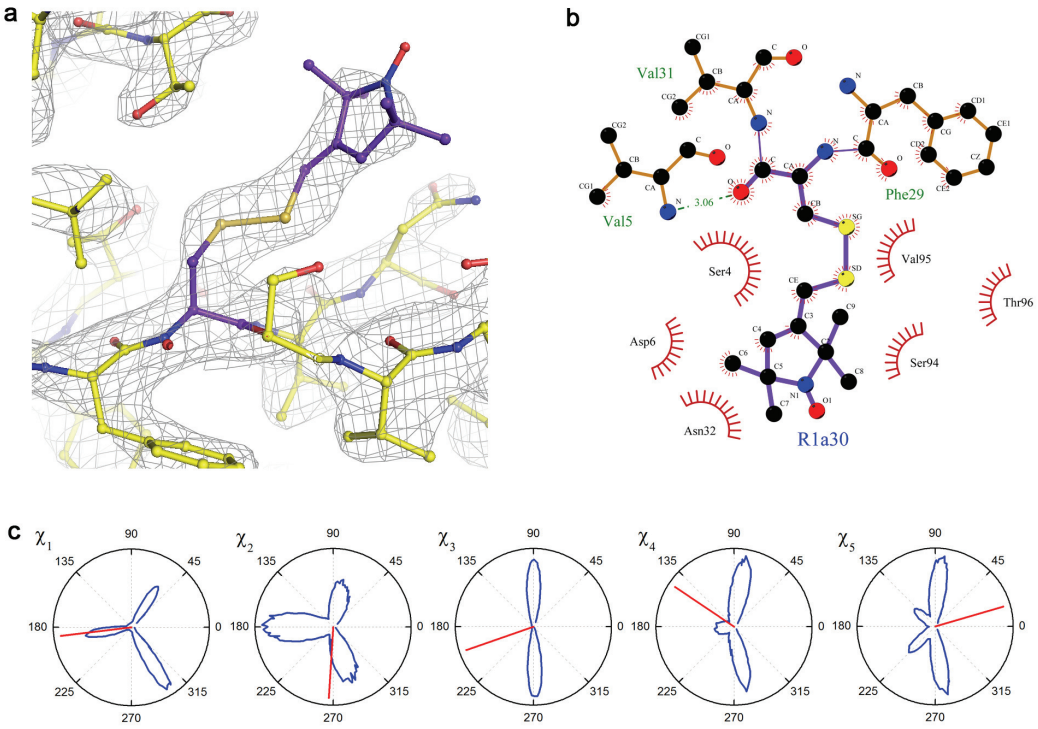


Fig. S5 continued: (caption see below)

T30R1, crystal II, chain M (B-factor = 79.3 Å^2)



T30R1, crystal II, chain N (B-factor = 71.4 Å^2)

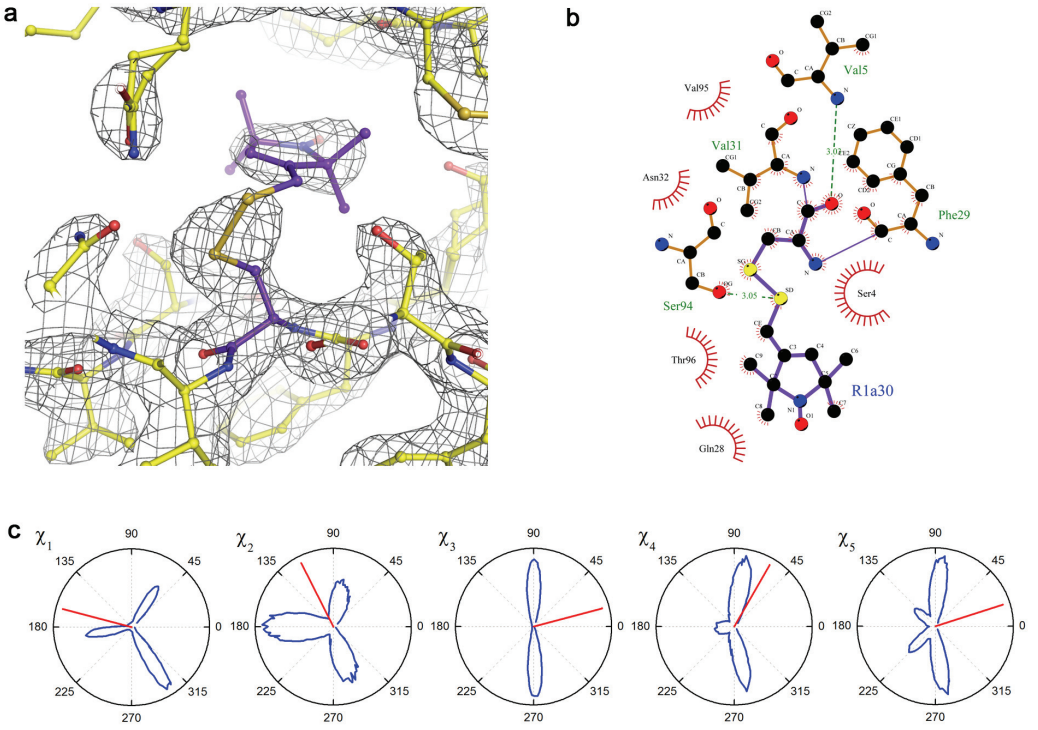


Fig. S5 continued: (caption see below)

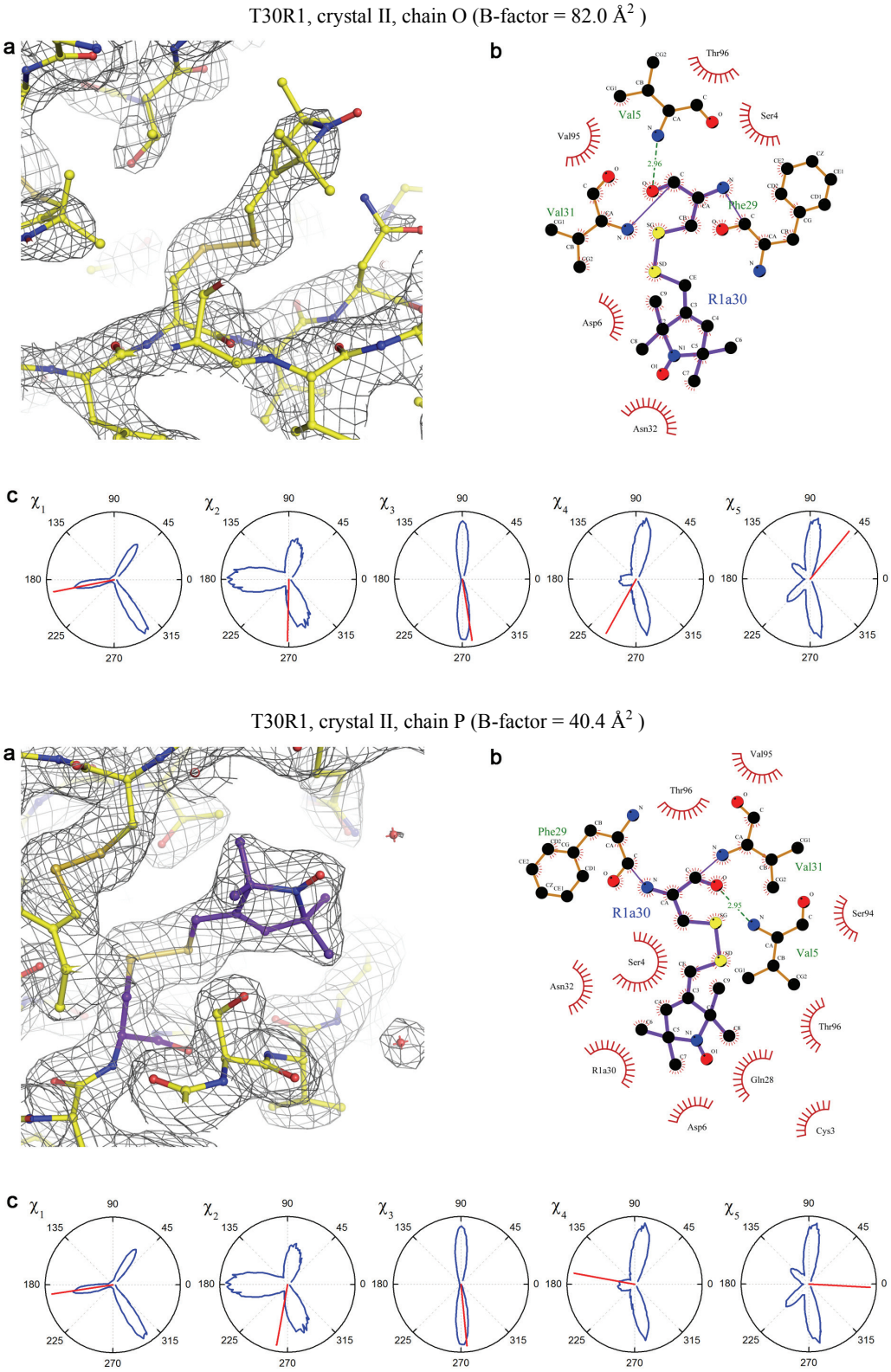


Fig. S5 Twenty X-ray conformers of the R1 side chain in T30R1. **(a)** Electron density of the R1 side chain. The protein is shown as a ball and stick model. The R1 side chain is highlighted in purple. The refined 2Fo-Fc electron density is shown as a grey mesh contoured at 1.0σ. The average B-factor of the side chain is given. Because only a single side chain was modelled for each individual monomer, the occupancy was set to 1.0 in each case. **(b)** Interaction topology diagram. Hydrophobic interactions are indicated as red arcs, polar interactions as dashed lines. Covalent bonds are shown as thick lines. **(c)** The dihedral angles χ_1- χ_5 that are determined for each X-ray conformer compared (red line) to the corresponding distributions from the MMM rotamer library (blue shapes) [Y. Polyhach, E. Bordignon and G. Jeschke, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2356].

Table S3. Dihedral angles of the R1 side chain in crystal structures of the azurin mutant T30R1.

Mutant, crystal form, chain	Quality of electron density[a]	χ_1 [°] [b]	χ_2 [°] [b]	χ_3 [°] [b]	χ_4 [°] [b]	χ_5 [°] [b]	Rotamer[c]
T21R1, I, A	+	-176	78	95	-60	88	*t(p)pmp*
T21R1, I, B	+	-176	73	95	-53	88	*t(p)pmp*
T21R1, I, C	+	-174	-143	93	169	101	*t(t)ptp*
T21R1, I, D	+	180	-140	100	175	96	*t(t)ptp*
T30R1, I, A	+	180	-94	-89	-59	86	*t(m)mmp*
T30R1, I, B	+	-176	-90	-85	-78	-17	*tmmm(m)*
T30R1, I, C	-	(143)	(-74)	(-86)	(-67)	(-155)	
T30R1, I, D	0	-151	-82	-80	(-80)	(-43)	*tmm*
T30R1, II, A	+	-178	-92	-86	171	-6	*t(m)mt(m)*
T30R1, II, B	+	-173	-99	-87	167	2	*t(m)mt(p)*
T30R1, II, C	+	-173	-97	-89	169	5	*t(m)mt(p)*
T30R1, II, D	-	175	(-136)	(-82)	(79)	(50)	*t*
T30R1, II, E	+	-176	-92	-85	159	5	*t(m)mt(p)*
T30R1, II, F	+	-173	-101	-85	-148	-81	*t(m)m(t)m*
T30R1, II, G	0	168	173	(73)	(-164)	(-5)	*tt*
T30R1, II, H	0	64	161	83	46	49	*ptpp(p)*
T30R1, II, I	0	93	178	114	-87	-32	*(p)tpm(m)*
T30R1, II, J	-	-168	-160	(-99)	(152)	(-1)	*tt*
T30R1, II, K	+	-176	-95	-82	162	2	*t(m)mt(p)*
T30R1, II, L	+	-170	-103	-83	-148	-87	*t(m)m(t)m*
T30R1, II, M	0	-173	-93	(-161)	(146)	(16)	*t(m)*
T30R1, II, N	0	165	117	(15)	(60)	(18)	*t(p)*
T30R1, II, O	0	-169	-91	-81	-119	51	*t(m)m(m)p*
T30R1, II, P	+	-171	-100	-84	170	-3	*t(m)mt(m)*

[a] The quality of electron density is classified as: "+" meaning all atoms are clearly visible; "0" meaning slight disorder but overall conformation is clear; "-" meaning the side chain is disordered.

[b] The values in brackets denote the angles that cannot be precisely determined from the electron density.

[c] The *t,p,m* notation of the rotamer states: for the angles χ_1, χ_2, χ_4 – p = 60°±30°, m = 300±30°, t = 180°±30°; for the angles χ_3, χ_5 – p = 90°±30°, m = -90±30° [Z. Guo, D. Cascio, K. Hideg K, T. Kálái and W. L. Hubbell, *Protein Sci.*, 2008, **17**, 228; G. Jeschke, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2013, **72**, 42]. The values in brackets denote the angles falling outside the ±30° interval. In these cases, the closest state is indicated.

[P6] Comparison of PELDOR and RIDME for distance measurements between nitroxides and low-spin Fe(III) ions

Reprinted with permission from

Dinar Abdullin, Fraser Duthie, Andreas Meyer, Elisa Müller, Gregor Hagelueken, Olav Schiemann, "Comparison of PELDOR and RIDME for Distance Measurements between Nitroxides and Low-Spin Fe(III) Ions", *J. Phys. Chem. B* **2015**, *119*, 13534-13542.
DOI: 10.1021/acs.jpcb.5b02118

Own contribution to the manuscript:

- Performing CD, CW-EPR, PELDOR and RIDME measurements
- Analysis of the PELDOR and RIDME data, comparing these methods
- Writing the manuscript

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

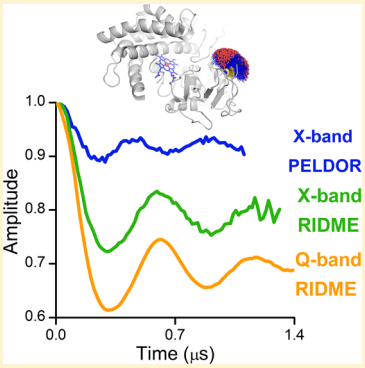
Comparison of PELDOR and RIDME for Distance Measurements between Nitroxides and Low-Spin Fe(III) Ions

Dinar Abdullin, Fraser Duthie, Andreas Meyer, Elisa S. Müller, Gregor Hagelueken, and Olav Schiemann*

Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstrasse 12, 53115 Bonn, Germany

Supporting Information

ABSTRACT: EPR-based nanometre distance measurements are becoming ever more important in structural biology. Usually the distance constraints are measured between two nitroxide spin labels. Yet, distance measurements between a metal center and spin labels enable, e.g., the localization of metal ions within the tertiary fold of biomolecules. Therefore, it is important to find methods that provide such distance information quickly, with high precision and reliability. In the present study, two methods, pulsed electron−electron double resonance (PELDOR) and relaxation-induced dipolar modulation enhancement (RIDME), are compared on the heme-containing and spin-labeled cytochrome P450cam. Special emphasis is put on the optimization of the dead-time free RIDME experiment and several ways of data analysis. It turned out that RIDME appears to be better suited for distance measurements involving metal ions like low-spin Fe^{3+} than PELDOR.



INTRODUCTION

Distance measurements by pulsed electron paramagnetic resonance (EPR) provide valuable information about the structure and dynamics of proteins and oligonucleotides.[1−3] The most commonly used technique is pulsed electron−electron double resonance (PELDOR).[4,5] It is often applied to measure the distance between two nitroxide spin labels site-specifically attached to a biomolecule.[2] In addition to nitroxides, there is a keen interest of using naturally occurring paramagnetic cofactors, like metal ions,[6−8] amino acids[9,10] or flavins,[11,12] for such distance measurements. In this way, new applications like the localization of metal ions within the fold of metalloproteins become possible.[13] So far, the majority of reported PELDOR studies on metal centers are dealing with low-spin Cu^{2+} ions,[6−8] iron−sulfur clusters[14,15] or manganese clusters.[16] Besides them, Gd^{3+} ions are often used as artificially introduced spin labels for PELDOR measurements.[17,18] All these centers usually display a moderate *g*-factor anisotropy and relatively long relaxation times. In contrast, PELDOR experiments involving metal ions, like Fe^{3+} or Mn^{2+}, are limited to a few recent reports.[19−22] This may stem from the fact that PELDOR measurements on such metal ions are usually time-consuming and demanding to analyze due to short relaxation times, low signal-to-noise ratio (SNR) and orientation selectivity. Furthermore, spin couples that consist of a nitroxide spin label and a metal center with largely different *g*-values than the nitroxide are difficult to study with the double-frequency PELDOR technique, due to the limited bandwidths of EPR resonators and microwave amplifiers. In this case, the use of single-frequency EPR techniques could be advantageous. One possibility, used already early on, would be to study how the relaxation of a slowly relaxing spin center, e.g., a nitroxide, changes due to the dipolar coupling to a faster relaxing spin center, e.g., a metal ion.[23] However, the extraction of distances by this relaxation method has a significant complexity that stems from the fact that the dipolar coupling leads only to an additional exponential decay of the signal and not to an oscillation and it needs the reference measurement without the metal center. In contrast, another single-frequency technique called relaxation-induced dipolar modulation enhancement (RIDME)[24] has been shown to yield time traces in which the dipolar coupling manifests itself as an oscillation of the signal. Therefore, we set up to compare RIDME and PELDOR, when applied to the spin couple low-spin Fe(III)/nitroxide. The advantage of the RIDME experiment being that the metal ion spin is flipped by spontaneous relaxation rather than by a microwave pulse as it is in PELDOR.

Since the pioneering work of Kulik et al.,[24] the RIDME experiment has been significantly advanced. The disadvantage of the original three-pulse RIDME sequence is its dead time, which obscures the initial part of the signal and, therefore, complicates the extraction of the distance distribution. Introduction of the four-pulse sequence allowed one to reduce the dead time,[25] and the five-pulse sequence recently proposed by the lab of Huber has been shown to be dead time free.[26] Although the three and four-pulse versions of RIDME were applied to a wide range of spin systems,[27−32] the use of the new dead-time free RIDME accounts only for a few studies on nitroxide/nitroxide and Gd^{3+}/Gd^{3+} model systems,[26,33] and on the Fe^{3+}−nitroxide spin couple in cytochrome *f*.[26] Note that the

Special Issue: Wolfgang Lubitz Festschrift

Received: March 4, 2015
Revised: May 20, 2015
Published: May 22, 2015

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

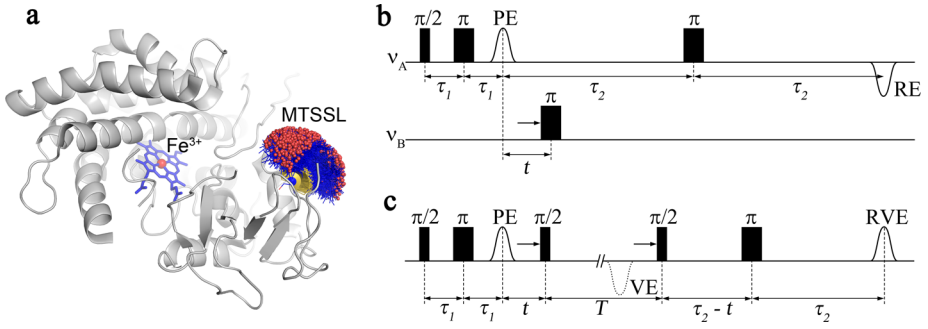


Figure 1. (a) Model of MTSSL-labeled cytochrome P450cam mutant C58R_1 created by means of the program mtsslWizard[35,36] using the crystal structure of the native form of the protein (PDB 3L61). The Fe^{3+}-MTSSL distances were measured using (b) the PELDOR and (c) the five-pulse RIDME experiments. The positions of the primary (PE), virtual (VE), refocused (RE), and refocused virtual (RVE) echoes are marked in the pulse sequences.

RIDME signals reported for the Fe^{3+}−nitroxide spin couple were shorter than 1 μs and showed no modulation, leaving the question of precise distance measurements for the Fe^{3+} spin center open. Another important aspect of the RIDME experiment concerns the removal of spin echo envelope modulations (ESEEM)[34] from the original signal. To do this, several approaches were introduced but have not been compared to each other.[26−29]

Therefore, we compare here PELDOR and the dead-time free RIDME experiments using the Fe^{3+}-containing protein cytochrome P450cam (CYP101) spin labeled at position C58 as a test case (Figure 1a). In addition, several methods for the suppression of ESEEM artifacts in the RIDME time trace are compared to each other.

MATERIALS AND METHODS

Cloning, Protein Expression, Purification, and Spin Labeling. The cytochrome P450cam gene from *Pseudomonas putida* was obtained via the Addgene repository in form of the pUS200 plasmid (https://www.addgene.org/17788/). The gene was PCR amplified using the following primers:

P450cam_fwd	TAATAACCATGGGGATGACGACTGAAACCATAC
P450cam_rev	TAATAACTCGAGTTATACCGCTTTGGTAGTCGC

The PCR product was digested with NcoI and XhoI and cloned into the pEHisTEV vector (Huanting Liu, University of St Andrews, U.K.). Native cysteines except C58 were removed by PCR techniques[37,38] using the following oligos:

C85S_fwd	GCGAGAGCCCGTTCATCCCTCGTGAAGCCG
C85S_rev	AACGGGCTCTCGCTGGAAAAGTGGCGGTAATCT
C136S_fwd	GCCAGCTCGCTGATCGAGAGCCTGCGCCCG
C136S_rev	GATCAGCGAGCTGGCCAGCTCCTGGATCCGGTT
C285S_fwd	CCGCTAGCGAGGAACTACTCCGGCGCTTCT
C285S_rev	TCCTCGCTAGCGGCTGGAATACGCTCGGGA
C334A_fwd	ACGCCGCTCCGATGCACGTCGACTTCAGTC
C334A_rev	ATCGGAGCGGCGTTTTCGCGCTCATCCAGG

In this way, the native cysteine at position 58 could be used for the spin labeling experiments. To express the protein, the plasmid was transformed into *Escherichia coli* C43 cells. The cells were grown at 37 °C in LB medium, supplemented with 50 μg/L kanamycin until an OD of 0.5 was reached. The cultures were then induced with 0.4 mM IPTG. Further, 2 mM 5-aminoleuvinic acid was added to increase the heme incorporation efficiency. After induction, the cells were incubated at 25 °C with shaking at 200 rpm and the expression was allowed to proceed overnight. After expression, the cells were harvested, resuspended in buffer A (50 mM sodium phosphate pH 6.0, 50 mM NaCl), and lysed using a constant systems cell disruptor. Cell debris and insoluble proteins were removed by centrifuging the lysate at 20 000 rpm for 20 min at 4 °C. The soluble fraction was loaded onto a 5 mL HisTrap Ni FF column (GE Healthcare), and the target protein was eluted using a gradient from 0 to 100% buffer B (50 mM sodium phosphate pH 6.0, 50 mM NaCl, 1 M imidazole). Peak fractions were pooled and dialyzed against buffer C (50 mM Tris pH 7.4, 300 mM NaCl) for 3 h at 4 °C before addition of 4 mg of TEV protease overnight at 4 °C. On the next day, the protease and uncleaved protein were removed by passing the sample through a 5 mL HisTrap Ni FF column. The flowthrough was concentrated using a Vivaspin 20 10.000 MWCO concentrator and loaded onto a Superdex 200 16/60 column (GE Healthcare). For labeling, a 5-fold excess of (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methanethiosulfonate (MTSSL) was added to the protein before incubation on ice for 3 h. Unbound label was later removed by means of a PD-10 column using either H_2O-based buffer (50 mM TES pH 7.4, 400 mM KCl) or D_2O-based buffer (50 mM TES pH 7.4, 400 mM KCl). Two samples obtained by this procedure were denoted in accordance to their buffer as C58R_1-H_2O and C58R_1-D_2O. For cryoprotection, 50% v/v ethylene glycol was added to C58R_1-H_2O and the same amount of deuterated ethylene glycol was added to C58R_1-D_2O. The final protein concentration in both samples was 150−200 μM.

EPR Measurements. All EPR measurements were carried out on a Bruker ELEXSYS E580 spectrometer. A Flexline probehead housing either an X-band (Bruker, ER 4118X-MD-5W1) or Q-band (Bruker, EN 5107D2) resonator was used. At the X-band all microwave pulses were amplified via a 1 kW TWT amplifier (Applied Systems Engineering, 117X). A standard Bruker 3W amplifier was used at the Q-band. To obtain low temperatures, a continuous flow helium cryostat (Oxford Instruments, CF935) and a temperature control system (Oxford Instruments, ITC 503S) were employed.

PELDOR experiments were performed with the standard four-pulse sequence shown in Figure 1b. The frequency of the pump pulse (ν_B) and the magnetic field were adjusted to excite the spins corresponding to the maximum of the nitroxide spectrum. The frequency of the detection pulses (ν_A) was set 250 MHz lower than the pump frequency, exciting a part of the

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

Fe^{3+} spins. The $\pi/2$ and π pulses of the detection sequence had lengths of 16 and 32 ns, and the pump pulse was 18 ns long. The $\pi/2$ pulse was phase-cycled to eliminate receiver offsets. The τ_1 and τ_2 intervals were set to 360 ns and 1.3 μs, respectively, which allowed one to acquire 90 data points with a 16 ns time increment. The PELDOR signal was recorded at 10 K with a repetition time of 1 ms. To achieve an acceptable SNR, the signal was acquired for 24 h.

RIDME experiments were performed with the five-pulse sequence shown in Figure 1c. Before setting up the pulse sequence, the value of the magnetic field was adjusted to set the frequency of the microwave pulses in resonance with the maximum of the nitroxide spectrum. Depending on the protein buffer and frequency band, different lengths of the pulses and interpulse intervals were used; all their values are listed in Table 1. To eliminate the contribution of unwanted echoes, an eight-step phase cycling[26] was applied. The time increment for the displacement of the third and fourth pulses, defining the time resolution of the RIDME signals, was 16 ns. Totals of 100 and 156 data points were acquired for the $C58R_1$ in H_2O and $C58R_1$ in D_2O samples, respectively. All experiments were performed at 25 K with a repetition time of 10 ms and 1000 averages per data point resulting in an overall measurement time of ~2.3 h for $C58R_1$-H_2O and ~3.5 h for $C58R_1$-D_2O.

Table 1. Parameters of the RIDME Experiments on the $C58R_1$ Samples

sample	microwave frequency (GHz)	$t_{\pi/2}$, t_π (ns)	τ_1 (ns)	τ_2 (μs)	T (μs)
$C58R_1$-H_2O	9.720	16, 32	140	1.6	30
$C58R_1$-H_2O	9.720	80, 160	246	1.6	30
$C58R_1$-H_2O	33.760	24, 48	140	1.6	30
$C58R_1$-D_2O	33.738	24, 48	140	2.5	30
$C58R_1$-D_2O	33.736	120, 240	340	2.7	30

The distance distributions were extracted from the RIDME and PELDOR time traces by means of the program DeerAnalysis[39] using a regularization parameter of 1. A second-order polynomial was empirically chosen as a background function and fitted for both PELDOR and RIDME time traces.

RESULTS AND DISCUSSION

Characterization of Samples. Prior to the distance measurements, the synthesized $C58R_1$ samples were characterized with respect to the protein identity and purity, the degree of spin labeling, flexibility of the spin label, the spin state and the content of Fe^{3+} ions. According to the typical reddish color, circular dichroism (CD) spectroscopy, and gel chromatography, the mutant does have the expected structure and is pure (Figure S1, Supporting Information). On the basis of the continuous wave (cw) X-band EPR spectra recorded at 293 K (Figure S2, Supporting Information), the labeling degree is estimated to be higher than 80%. The cw EPR spectra also reveal that the protein-bound labels are significantly immobilized and adopt at least two distinct conformations (Supporting Information). The low-temperature cw EPR spectra showed the typical low-spin heme signal for cytochrome P450cam with the g-tensor (g_1, g_2, g_3) = (2.42, 2.25, 1.91) and no traces of the high-spin heme.[40] Using UV−vis spectroscopy, the ratio of heme per cytochrome molecule was roughly estimated to be 52% for the $C58R_1$-H_2O sample and 40% for the $C58R_1$-D_2O sample (Figure S3, Supporting Information). Note that the difference between the two samples is related to the fact that both samples were synthesized independently from each other.

Setting Up the Distance Measurements. Before performing the PELDOR and RIDME measurements on the Fe^{3+}-nitroxide spin pair one has to choose which spin will be used for recording the echo signal (referred to as spin A) and which spin will be flipped (referred to as spin B). The criteria of this assignment are different for PELDOR and RIDME. An efficient RIDME experiment requires the spin−lattice relaxation time of spin B to be much shorter than that of spin A.[24] This requirement is fulfilled, when spin A corresponds to a nitroxide and spin B is the faster relaxing Fe^{3+} center. In contrast, the use of the nitroxide as spin B is preferable for the PELDOR experiment, because this yields the larger modulation depth. Then, spin A will correspond to a subset of Fe^{3+} spins. Note that only those Fe^{3+} spins that have g-values close to g_3 could be used as spin A here, because the spectral separation between the nitroxide signal and the low-field part of the Fe^{3+} signal is much larger than the resonator bandwidth. The spectral positions of the microwave pulses are indicated in Figure 2 for the PELDOR and the RIDME setup.

PELDOR Measurements. As mentioned above, the PELDOR signal was acquired on the high-field region of the

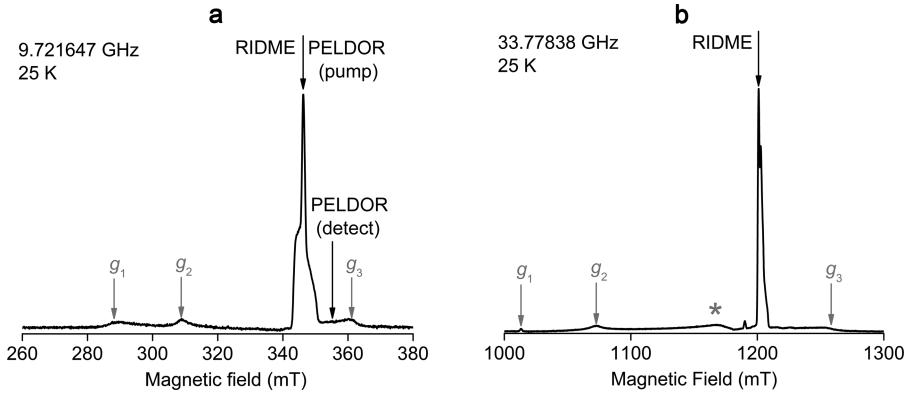


Figure 2. Field-swept echo-detected spectrum of $C58R_1$-H_2O recorded at (a) the X-band and (b) the Q-band. The components of the Fe^{3+} effective g-factors are shown by gray arrows. The spectral positions of the microwave pulses are indicated by black arrows. The signal in the Q-band spectrum marked by an asterisk corresponds to the copper background of the resonator.

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

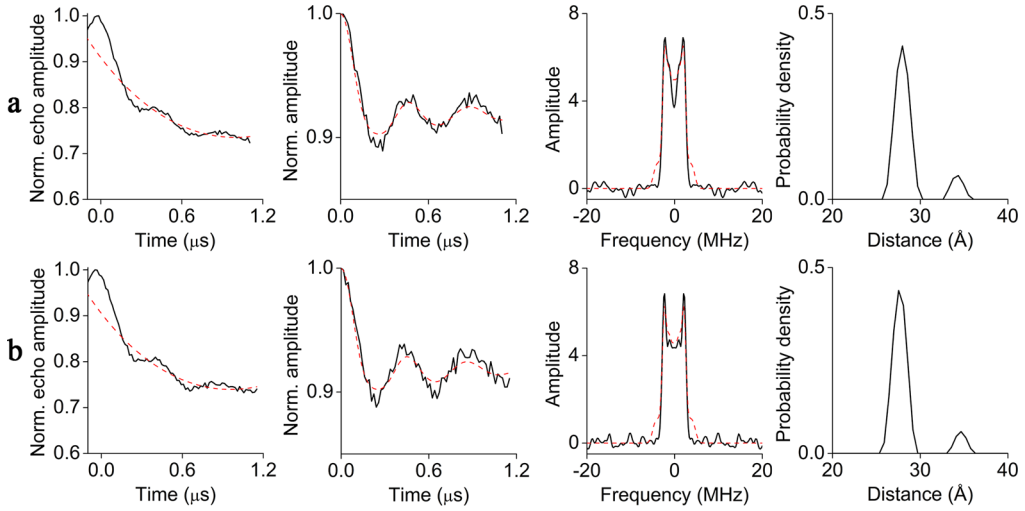


Figure 3. PELDOR data of (a) C58R$_1$-H$_2$O and (b) C58R$_1$-D$_2$O samples. Column 1: primary time traces (black lines) and their background fits (red dashed line). Column 2: background-subtracted time traces (black lines) and their fits by DeerAnalysis (red dashed lines). Column 3: FFT of the background-subtracted time traces and their fits by DeerAnalysis (red dashed lines). Column 4: derived Fe^{3+}-MTSSL distance distributions.

Fe^{3+} spectrum, whereas the nitroxide spins were flipped by the pump pulse. Because the phase memory time (T_m) of the Fe^{3+} ions in the C58R$_1$-H$_2$O and C58R$_1$-D$_2$O samples were found to be nearly the same (Figure S4, Supporting Information), the same interpulse intervals were used for both samples in the PELDOR experiments. The PELDOR time traces of both samples look identical (Figure 3). They have a modulation depth of 8% and comprise two periods of the dipolar modulation. The reasons for this weak modulation depth are incomplete labeling and orientation selectivity, which occurs because the detection pulses excite only a small fraction of the Fe^{3+} spectrum. The evidence of orientation selectivity can be clearly seen in the fast Fourier transform (FFT) of the time traces, which do not reproduce the shape of the Pake doublet (Figure 3). In the presence of orientation selectivity the extraction of distances by the program DeerAnalysis is not reliable anymore. For the accurate analysis of orientation selective PELDOR data, several time traces recorded on different spectral components of anisotropic spin centers are required.[41−43] This requirement could not be fulfilled here, because the separation between the nitroxide spectrum and the low-field components of the Fe^{3+} spectrum (g_1 and g_2 in Figure 2) is much larger than the resonator bandwidth and even the available frequency range of a commercial X-band spectrometer. Nevertheless, for the sake of comparison with the RIDME results, the obtained PELDOR time trace was analyzed by DeerAnalysis (Figure 3). A relatively good fit to the time trace was achieved, yielding a distance distribution with a mean value of 27.7 Å and a standard deviation of 0.8 Å. The small distance peak appearing at 35 Å is attributed to incomplete background removal and orientation selectivity.

RIDME Measurements. The five-pulse sequence illustrated in Figure 1c was used to perform the RIDME experiment. The Fe^{3+} spins, which were chosen to be the B spins, have to flip by spontaneous relaxation during the time interval T of this sequence. Thus, the value of T should be adjusted so that the majority of the Fe^{3+} spins are flipped. Because the relaxation rate of the Fe^{3+} spins is temperature dependent, the adjustment of T should be done together with the temperature adjustment to ensure the following additional requirements: T should not be too long, to minimize spectral diffusion effects, but still much longer than the maximal interval t, so that the spin−lattice relaxation of the Fe^{3+} spin during t can be neglected.[24] To determine T experimentally, a series of inversion recovery curves were recorded on the C58R$_1$-H$_2$O sample at different temperatures (Figure 4). A T value of 30 μs, which corresponds

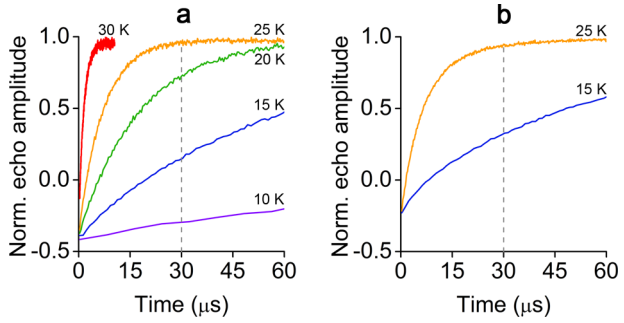


Figure 4. Inversion recovery for the Fe^{3+} spin center in C58R$_1$-H$_2$O recorded at different temperatures and at (a) the X-band and (b) the Q-band. The dashed line corresponds to the T value chosen for RIDME.

to almost complete recovery of the Fe^{3+} longitudinal magnetization at 25 K and, therefore, to the highest probability of the Fe^{3+} spin flip at this temperature, was chosen for the RIDME measurements.[29] Note that the chosen values of T and temperature provide a good compromise between the requirements mentioned above (Figures S5 and S6, Supporting Information).

The X-band RIDME time trace recorded on the C58R$_1$-H$_2$O sample is shown in Figure 5a. The FFT of this signal consists of the Pake pattern appearing at frequencies <5 MHz and a peak around 14.7 MHz corresponding to the ^{1}H Larmor frequency (Figure 5b). Although the dipolar and ESEEM frequencies are well separated from each other, it is worthwhile to get rid of the proton ESEEM for the simplicity of the further RIDME data analysis. To do this, four different approaches were considered.

As proposed in previous works,[29,30] ESEEM can be suppressed through dividing an original RIDME signal by a reference RIDME signal which contains ESEEM, but no dipolar modulation. Such a reference signal is obtained, when the

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

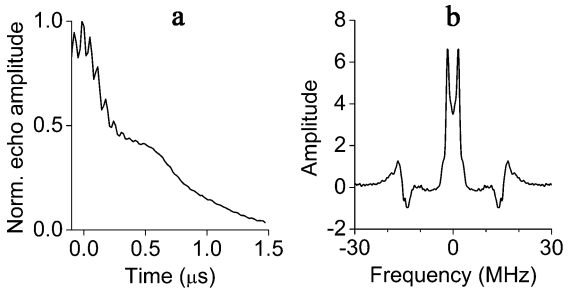


Figure 5. (a) RIDME time trace acquired on C58R$_1$-H$_2$O at the X-band and (b) its FFT.

inversion of the Fe^{3+} spins during the time interval T is made inefficient. The conditions leading to inefficient inversion are readily determined from Figure 4: either the T interval should be shortened, i.e., from 30 μs down to 2 μs, or the temperature of the experiment should be lowered, i.e., from 25 K down to 15 K. Both conditions were applied here to record the reference RIDME signal, which was then used to normalize the original RIDME signal. The resulting normalized time traces are compiled in Figure 6a,b. They reveal that the normalization via temperature allows one to remove ESEEM completely, whereas after the normalization via T some ESEEM artifacts are still present in the data.

Another possibility to suppress ESEEM is to use selective microwave pulses. Note that the RIDME signal described above was acquired with rather broadband $\pi/2$ and π pulses of 16 and 32 ns, respectively. Because the excitation bandwidths corresponding to these pulses (37.5 and 18.7 MHz) are larger than the ^{1}H Larmor frequency, proton ESEEM contributes to the signal. However, when the excitation bandwidth of the microwave pulses is much smaller than the ^{1}H Larmor frequency, it can be suppressed.[31] The RIDME signal acquired on the C58R$_1$-H$_2$O sample with 80 and 160 ns pulses is shown in Figure 6c. As can be seen, the signal contains the dipolar modulation and, importantly, no traces of proton ESEEM. Besides, this approach does not require the acquisition of a reference time trace reducing the measurement time by a factor

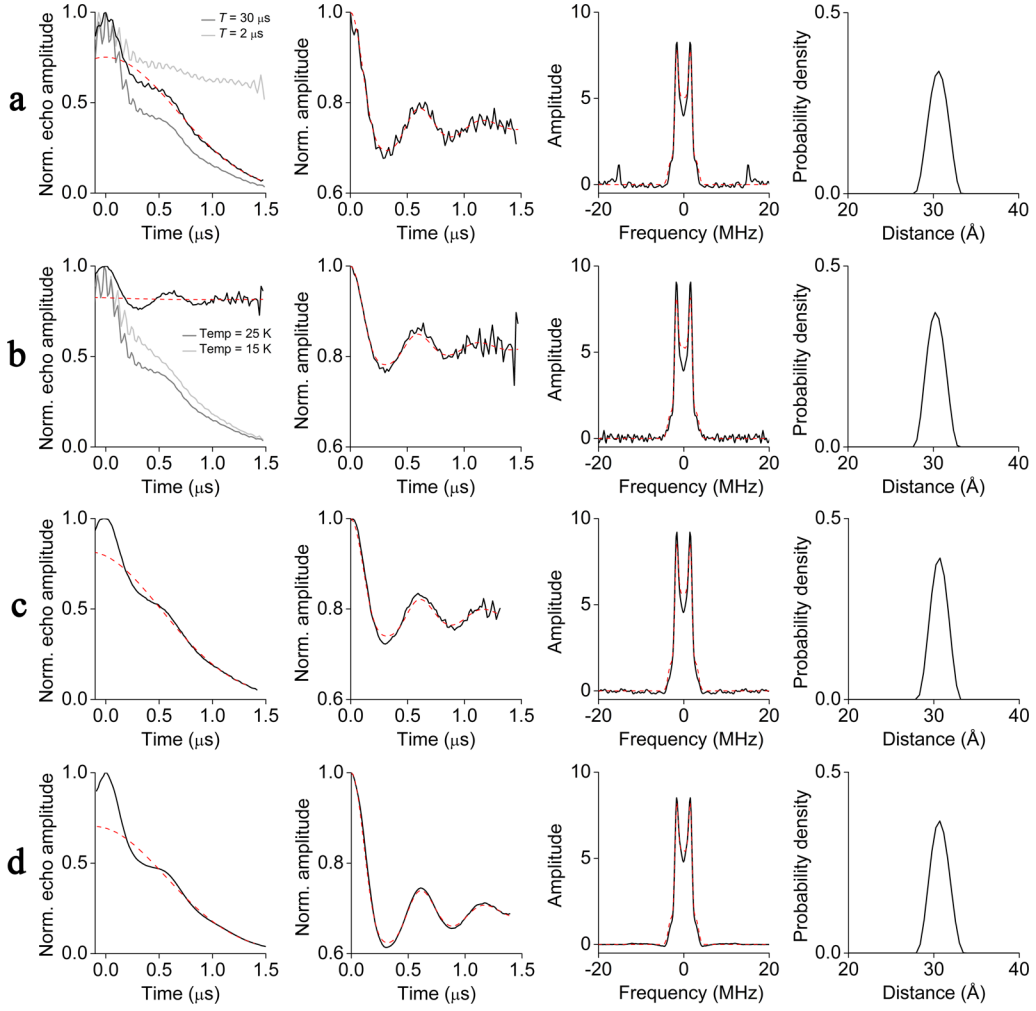


Figure 6. Different approaches of ESEEM removal from the RIDME signal acquired on C58R$_1$-H$_2$O: (a) ESEEM is removed by division of two RIDME signals recorded with T = 30 μs and T = 2 μs. The division result is shown as a black line. (b) ESEEM is removed by division of two RIDME signals recorded at 25 and 15 K. The division result is shown as a black line. (c) ESEEM is suppressed by using selective pulses with lengths of 80 and 160 ns for π and $\pi/2$ pulses, respectively. (d) ESEEM is suppressed by performing RIDME at Q-band frequencies. Column 1: primary time traces (black lines) and their background fits (red dashed lines). Column 2: background-subtracted time traces (black lines) and their fits by DeerAnalysis (red dashed lines). Column 3: FFT of the background-subtracted time traces and their fits by DeerAnalysis (red dashed lines). Column 4: derived Fe^{3+}-MTSSL distance distributions.

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

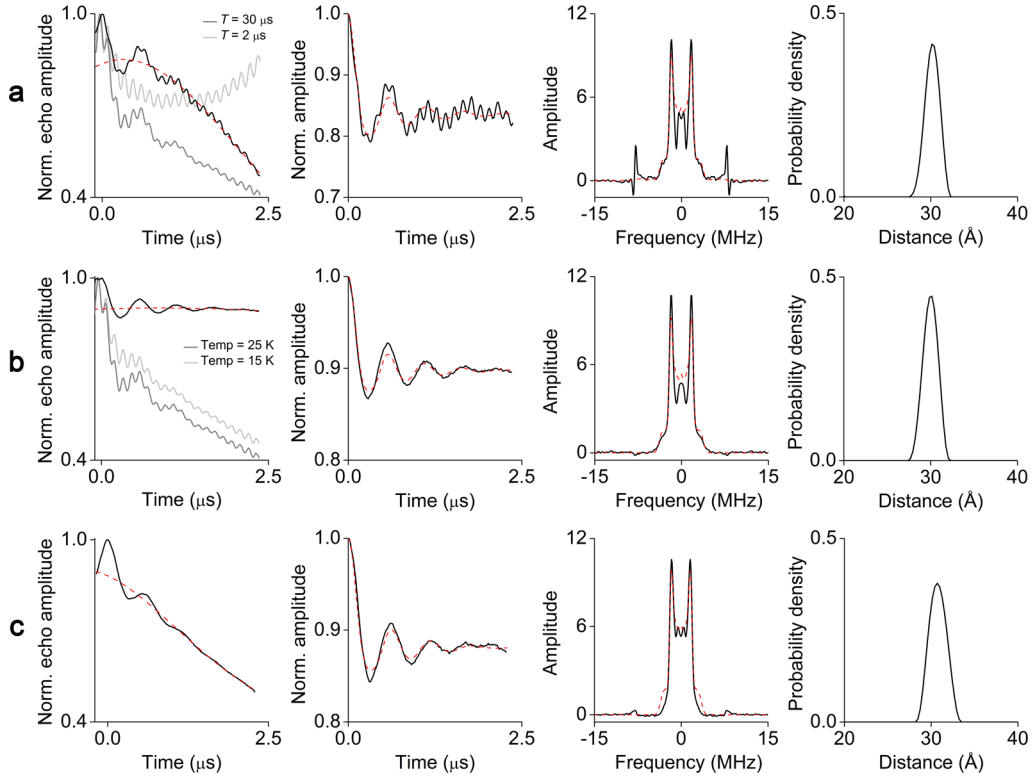


Figure 7. Different approaches of ESEEM removal from the Q-band RIDME signal acquired on C58R$_1$-D$_2$O. (a) ESEEM is removed by division of two RIDME signals recorded with T = 30 μs and T = 2 μs. The division result is shown by black line. (b) ESEEM is removed by division of two RIDME signals recorded at 25 and 15 K. The division result is shown by black line. (c) ESEEM is suppressed by using 120 and 240 ns pulses. Column 1: primary time traces (black lines) and their background fits (gray dashed lines). Column 2: background-subtracted time traces (black lines) and their fits by DeerAnalysis (gray dashed lines). Column 3: FFT of the background-subtracted time traces and their fits by DeerAnalysis (gray dashed lines). Column 4: derived Fe^{3+}-MTSSL distance distributions.

of 2. The same also implies the use of high microwave frequencies, which reduce the probability of forbidden transitions and, consequently, the amplitude of ESEEM.[34] Moreover, because the ^{1}H Larmor frequency increases proportionally with the microwave frequency/magnetic field ($\nu_I(^1H)$ = 51.1 MHz at 1.2 T), the condition when the excitation bandwidth of the microwave pulses is much smaller than the ^{1}H Larmor frequency can be achieved without significant elongation of the pulses. The RIDME signal recorded at the Q-band confirms the efficiency of this approach (Figure 6d).

The conversion of the RIDME time traces into distance distributions was done by means of the program DeerAnalysis. First, a decay with an empirically chosen shape of second-order polynomial was removed from the time traces (Figure 6, column 1). As mentioned in previous work,[33] a theoretical description of the RIDME background is still absent, but one important contribution to it is probably spin diffusion caused by interactions of the nitroxide spin with surrounding magnetic nuclei.[26] However, because the time traces exhibit a prominent modulation, the influence of the background function on the subsequently derived distances is expected to be negligible. The background-corrected time traces feature modulation depths of 18−30% (Figure 6, column 2). This variation of the modulation depth may be caused by differences in the ESEEM suppression procedures and the error of the background fitting. Note that the division of the two time traces used to remove the ESEEM reduces the contribution of the background to the RIDME signals but does not eliminate it completely (Figure 6a,b). The obtained modulation depths are lower than the theoretically achievable value of 50%,[26] which is attributed to the fraction of cytochrome molecules containing no heme group as described above and to free nitroxides. In principle, the low modulation depth could also stem from a nonoptimal value for T. However, this possibility was excluded by recording the dependence of the modulation depth parameter on T (Figure S5, Supporting Information). The time traces corresponding to different procedures of ESEEM suppression have also slightly different SNRs. The lowest average SNR corresponds to the procedures using a reference time trace. This is probably due to the division of two low-amplitude signals at high t values. The time trace recorded with selective pulses might have been expected to show the lowest SNR, because fewer spins are excited, but because it does not require the division by a reference signal, it actually has a SNR that is ∼5 times higher. Also the Q-band time trace has a gain in SNR of ∼68 compared to the reference procedures at the X-band. FFT of the time traces yield the characteristic Pake doublet, meaning that the RIDME experiment recorded on the maximum of the nitroxide spectrum has apparently no orientation selectivity (Figure 6, column 3). In contrast, recording the RIDME experiment at other positions on the nitroxide spectrum shows that orientation selection can be introduced into RIDME if wanted (Figure S7, Supporting Information). Finally, the time traces were transformed into a distance distributions using Tikhonov regularization. A good fit was obtained for the distance distribution with a mean value of 30.5 Å and a standard deviation of 1.0 Å. Significantly, the distance distributions are invariant with respect to the different

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

ESEEM suppression procedures (Figure S8, Supporting Information).

Due to the limited T_m of the nitroxide spins, the RIDME signals decay to zero after roughly 1.5 μs (Figure 6). Nevertheless, this time window allows the robust determination of distances up to 45 Å.[2] To increase this limit, T_m of the nitroxide spins was prolonged by using deuterated buffer for the C58R$_1$-D$_2$O sample.[44] Note that the ^{2}D Larmor frequency is similar to the observed dipolar frequencies at the X-band ($\nu_I(^2D)$ = 2.29 MHz at 350 mT) and bigger than these frequencies at the Q-band ($\nu_I(^2D)$ = 7.85 MHz at 1.2 T). Therefore, to avoid the complexity related to the deuterium ESEEM removal at the X-band, the RIDME experiments were performed at the Q-band. Running the Q-band RIDME experiment with the same pulse lengths as for the C58R$_1$-H$_2$O sample (24/48 ns) yielded time traces that still contain a considerable ESEEM contribution (Figure 7a,b). Thus, to eliminate the ESEEM artifacts, the Q-band RIDME signal was normalized with a reference signal. Similar to the C58R$_1$-H$_2$O sample, normalization of the RIDME time trace via temperature is more efficient than normalization via T.

In the following, the time trace obtained after normalization via temperature is discussed exemplarily (Figure 7b). Due to the deuterated buffer, a nonzero echo signal is detected even after 2.5 μs, which allows observing more than two well-resolved periods of the dipolar modulation. However, the obtained modulation depth of 10% is lower than the modulation depths observed for the C58R$_1$-H$_2$O sample. This is attributed to the lower heme content in the C58R$_1$-D$_2$O (40%) as compared to the C58R$_1$-H$_2$O (52%) sample. The distance distribution obtained from this time trace yields a mean distance of 30.1 Å and a standard deviation of 0.8 Å, which is in agreement with the distances found for C58R$_1$-H$_2$O.

The use of selective 120 and 240 ns pulses not only led to nearly complete suppression of ESEEM but also caused the loss of the high-frequency component of the Pake pattern (Figure 7c). The lack of this component may either be due to orientation selectivity resulting from the selective excitation of nitroxide spins by the long microwave pulses or may be due to the suppression of the largest dipolar frequencies by the microwave pulses in the same way as they suppress ESEEM. The second assumption is supported by the fact that the excitation bandwidth corresponding to 240 ns π-pulse (2.5 MHz) is lower than the missing frequencies of the Pake doublet.

Comparison of the PELDOR and RIDME Results. Measurements of Fe^{3+}−MTSSL distances conducted on the C58R$_1$ samples reveal several advantages of RIDME over PELDOR. First of all, the RIDME experiment allows obtaining the whole dipolar spectrum (Pake doublet), whereas some dipolar frequencies are absent in the PELDOR-derived spectrum due to orientation selectivity. This makes the analysis of RIDME data easier and the corresponding distance distribution more reliable.

Second, the RIDME time traces have better SNR and therefore require shorter measurement time as compared to the PELDOR time trace. The X-band RIDME signal (Figure 6c) acquired on the C58R$_1$-H$_2$O sample has a 7 times higher SNR as compared to the corresponding X-band PELDOR signal (Figure 3a). In addition, to account for orientation selectivity in PELDOR data analysis, several time traces at different orientations would have to be recorded that would increase the measurement time, thus favoring the RIDME experiment even more. Recording the PELDOR time traces at the Q-band with a 150 W TWT amplifier may increase SNR by a factor of 13−20,[45] but a similar increase in SNR is then expected also for the RIDME time trace (Figure 6c,d). Moreover, increased orientation selectivity is expected for Q-band PELDOR time traces, because the Fe^{3+} spectrum is g-dominated and therefore broader at Q-band than at X-band.

Importantly, the use of deuterated buffer allowed one to increase the length of the RIDME time trace up to 2.5 μs. At the same time, the length of the PELDOR time trace reaches 1.5 μs only, because the signal is recorded on the fast relaxing Fe^{3+} centers. In principle, the PELDOR time trace could be recorded on the slower relaxing nitroxide spin, while flipping the Fe^{3+} spins. However, this would lead to a small modulation depth, due to the small part of the Fe^{3+} spins flipped by the pump pulse, and longer shot repetition time, because of the slower longitudinal relaxation of the nitroxide.

The modulation depths of the PELDOR and RIDME time traces are with ∼10% similar for the C58R$_1$-D$_2$O sample and differ by 10−15% for the C58R$_1$-H$_2$O sample. Note that each method has its own specific factors influencing the value of the modulation depth. As mentioned above, the RIDME modulation depth is affected by the heme content in the samples. At the same time, the PELDOR modulation depth is dependent here on the labeling degree and orientation selection.

Both methods require the suppression of ESEEM artifacts in their time traces. In PELDOR data these artifacts are weak and can be suppressed by recording time traces for the several τ values. In contrast, the RIDME experiment can have considerable ESEEM contributions in the time trace. Nevertheless, four different ways of ESEEM removal applied for RIDME have been shown to be efficient and easy to implement.

The distance distributions obtained by PELDOR and RIDME are overlaid in Figure 8. The corresponding mean

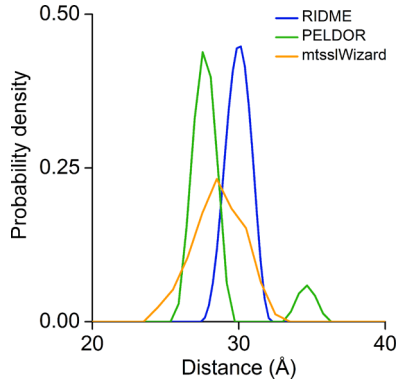


Figure 8. Comparison of the experimental and simulated Fe^{3+}-MTSSL distances in the mutant C58R$_1$.

distances differ by 2.4 Å. This difference is mainly attributed to the error of the PELDOR-derived distance, which stems from the inexact analysis of orientation selective PELDOR data as well as from the short length of the PELDOR time trace. Because of the same reasons, the similarity of the distribution widths is considered as accidental. Nevertheless, both experimental distributions fit to the prediction, obtained from molecular modeling by the program mtsslWizard[35,36] and using the crystal structure of cytochrome P450cam (PDB 3L61). In

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

addition, the narrow distance distribution observed by RIDME can be correlated to the significantly reduced flexibility of the label revealed by the cw X-band EPR spectra (see above).

CONCLUSION

The five-pulse RIDME experiment has been shown to be more suitable than PELDOR for the distance measurements between the Fe^{3+} ion and nitroxide spin label. The main advantages of the RIDME experiment are that it is less affected by orientation selectivity, which can even be avoided in the present case, and that the SNR is 7 times higher. This, in turn, allows for an easy and robust conversion of the RIDME signals into distance distributions. The use of deuterated protein buffer was shown to provide a longer RIDME signal but may, for the sake of separating dipolar frequencies from ESEEM artifacts, require the use of microwave frequencies higher than X-band.

Four different methods of ESEEM suppression in the RIDME signal were tested. The method in which the original signal is normalized by the reference signal recorded at lower temperature was shown to be more efficient than the one in which the reference signal was obtained by varying the T interval. The use of Q-band frequency allowed the suppression of proton ESEEM but was not high enough to suppress deuterium ESEEM. As an alternative method, the use of long selective microwave pulses was successfully applied. However, it requires the dipolar frequencies to be lower than the ESEEM frequencies that will be not always fulfilled for short distances at particular microwave frequency. Importantly, we have shown that none of these methods for ESEEM removal affected the final distance distribution. Thus, five-pulse RIDME holds great promise for distance measurements to metal centers. In addition, it was shown recently that five-pulse RIDME can be applied to high-spin metal ions although higher harmonics of the dipolar coupling constant have to be taken into account.[33]

Notably, the use of arbitrary waveform broadband microwave pulses may improve the performance of PELDOR measurements for the Fe^{3+}−nitroxide spin pairs with respect to SNR, orientation selectivity and modulation depth as recently shown for bisnitroxides.[46] If one can excite the whole nitroxide spectrum with the broadband pulses, RIDME measurements on given spin system may also get advantage from using such pulses.

ASSOCIATED CONTENT

Supporting Information

The SDS-PAGE analysis, size-exclusion chromatogram, CD spectra, cw EPR spectra, UV−vis spectra, 2-pulse ESEEM of the cytochrome P450cam mutant C58R_1. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b02118.

AUTHOR INFORMATION

Corresponding Author

*O. Schiemann. E-mail: schiemann@pc.uni-bonn.de. Phone: +49(0)228-732989.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support of the Deutsche Forschungsgemeinschaft through the Collaborative Research Center SFB813 “Chemistry at Spin Centers”. We also thank Prof. Arne Lützen and Christina Tenten (University of Bonn) for assistance with CD measurements.

REFERENCES

(1) Schiemann, O.; Prisner, T. F. Long-Range Distance Determinations in Biomacromolecules by EPR Spectroscopy. *Q. Rev. Biophys.* **2007**, *40*, 1−53.
(2) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419−446.
(3) Borbat, P. P.; Freed, J. H. Pulse Dipolar Electron Spin Resonance: Distance Measurements. In *Structure and Bonding (Berlin)*; Timmel, C. R., Harmer, J., Eds.; Springer: Heidelberg, 2013; Vol. *152*, pp 206−210, pp 1−82.
(4) Milov, A. D.; Salikhov, K. M.; Shirov, M. D. Application of ELDOR in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. *Fiz. Tverd. Tela* **1981**, *23*, 975−982.
(5) Martin, R. E.; Pannier, M.; Diederich, F.; Gramlich, V.; Hubrich, M.; Spiess, H. W. Determination of End-to-End Distances in a Series of TEMPO Diradicals of up to 2.8 nm Length with a New Four-Pulse Double Electron Electron Resonance Experiment. *Angew. Chem., Int. Ed.* **1998**, *37*, 2833−2837.
(6) van Amsterdam, I. M. C.; Ubbink, M.; Canters, G. W.; Huber, M. Measurement of a Cu-Cu Distance of 26 Å by a Pulsed EPR Method. *Angew. Chem., Int. Ed.* **2003**, *42*, 62−64.
(7) Kay, C. W. M.; El Mkami, H.; Cammack, R.; Evans, R. W. Pulsed ELDOR Determination of the Intramolecular Distance between the Metal Binding Sites in Dicupric Human Serum Transferrin and Lactoferrin. *J. Am. Chem. Soc.* **2007**, *129*, 4868−4869.
(8) Yang, Z.; Kurpiewski, M. R.; Ji, M.; Townsend, J. E.; Mehta, P.; Jen-Jacobson, L.; Saxena, S. ESR Spectroscopy Identifies Inhibitory Cu^{2+} Sites in a DNA-Modifying Enzyme to Reveal Determinants of Catalytic Specificity. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E993−E1000.
(9) Denysenkov, V. P.; Prisner, T. F.; Stubbe, J.; Bennati, M. High-Field Pulsed Electron-Electron Double Resonance Spectroscopy to Determine the Orientation of the Tyrosyl Radicals in Ribonucleotide Reductase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13386.
(10) Denysenkov, V. P.; Biglino, D.; Lubitz, W.; Prisner, T. F.; Bennati, M. Structure of the Tyrosyl Biradical in Mouse R2 Ribonucleotide Reductase from High-Field PELDOR. *Angew. Chem., Int. Ed.* **2008**, *47*, 1224−1227.
(11) Kay, C. W. M.; Elsässer, C.; Bittl, R.; Farrell, S. R.; Thorpe, C. Determination of the Distance between the Two Neutral Flavin Radicals in Augmenter of Liver Regeneration by Pulsed ELDOR. *J. Am. Chem. Soc.* **2006**, *128*, 76−77.
(12) Swanson, M. A.; Kathirvelu, V.; Majtan, T.; Frerman, F. E.; Eaton, G. R.; Eaton, S. S. DEER Distance Measurement Between a Spin Label and a Native FAD Semiquinone in Electron Transfer Flavoprotein. *J. Am. Chem. Soc.* **2009**, *131*, 15978−15979.
(13) Abdullin, D.; Florin, N.; Hagelueken, G.; Schiemann, O. EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules. *Angew. Chem., Int. Ed.* **2015**, *54*, 1827−1831.
(14) Elsässer, C.; Brecht, M.; Bittl, R. Pulsed Electron-Electron Double Resonance on Multinuclear Metal Clusters: Assignment of Spin Projection Factors Based on the Dipolar Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 12606−12611.
(15) Roessler, M. M.; King, M. S.; Robinson, A. J.; Armstrong, F. A.; Harmer, J.; Hirst, J. Direct Assignment of EPR Spectra to Structurally Defined Iron-Sulfur Clusters in Complex I by Double Electron-Electron Resonance. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1930−1935.
(16) Kawamori, A.; Katsuta, N.; Mino, H.; Ishii, A.; Minagawa, J.; Ono, T.-A. Positions of Q_A and Chl_Z Relative to Tyrosine Y_Z and Y_D in Photosystem II Studied by Pulsed EPR. *J. Biol. Phys.* **2002**, *28*, 413−426.
(17) Goldfarb, D. Gd^{3+} Spin Labeling for Distance Measurements by Pulse EPR Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9685−9699.

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

(18) Yulikov, M. Spectroscopically Orthogonal Spin Labels and Distance Measurements in Biomolecules. *Electron Paramagn. Reson.* **2015**, *24*, 1−31.
(19) Ezhevskaya, M.; Bordignon, E.; Polyhach, Y.; Moens, L.; Dewilde, S.; Jeschke, G.; Van Doorslaer, S. Distance Determination between Low-Spin Ferric Haem and Nitroxide Spin Label Using DEER: the neuroglobin case. *Mol. Phys.* **2013**, *111*, 2855−2864.
(20) Banerjee, D.; Yagi, H.; Huber, T.; Otting, G.; Goldfarb, D. Nanometer-Range Distance Measurement in a Protein Using Mn^{2+} Tags. *J. Phys. Chem. Lett.* **2012**, *3*, 157−160.
(21) Akhmetzyanov, D.; Plackmeyer, J.; Endeward, B.; Denysenkov, V.; Prisner, T. F. Pulsed Electron-Electron Double Resonance Spectroscopy between a High-Spin Mn^{2+} Ion and a Nitroxide Spin Label. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6760−6766.
(22) Kaminker, I.; Bye, M.; Mendelman, N.; Gislason, K.; Sigurdsson, S. Th.; Goldfarb, D. Distance Measurements between Manganese(II) and Nitroxide Spin-Labels by DEER Determine a Binding Site of Mn^{2+} in the HP92 Loop of Ribosomal RNA. *Phys. Chem. Chem. Phys.* **2015**, DOI: 10.1039/C5CP01624J.
(23) Rakowsky, M. H.; More, K. M.; Kulikov, A. V.; Eaton, G. R.; Eaton, S. S. Time-Domain Electron Paramagnetic Resonance as a Probe of Electron-Electron Spin-Spin Interaction in Spin-Labeled Low-Spin Iron Porphyrins. *J. Am. Chem. Soc.* **1995**, *117*, 2049−2057.
(24) Kulik, L. V.; Dzuba, S. A.; Grigoryev, I. A.; Tsvetkov, Y. D. Electron Dipole-Dipole Interaction in ESEEM of Nitroxide Biradicals. *Chem. Phys. Lett.* **2001**, *343*, 315−324.
(25) Kulik, L. V.; Grishin, Y. A.; Dzuba, S. A.; Grigoryev, I. A.; Klyatskaya, S. V.; Vasilevsky, S. F.; Tsvetkov, Y. D. Electron Dipole-Dipole ESEEM in Field-Step ELDOR of Nitroxide Biradicals. *J. Magn. Reson.* **2002**, *157*, 61−68.
(26) Milikisyants, S.; Scarpelli, F.; Finiguerra, M. G.; Ubbink, M.; Huber, M. A Pulsed EPR Method to Determine Distances between Paramagnetic Centers with Strong Spectral Anisotropy and Radicals: the Dead-Time Free RIDME Sequence. *J. Magn. Reson.* **2009**, *201*, 48−56.
(27) Savitsky, A.; Dubinskii, A. A.; Flores, M.; Lubitz, W.; Möbius, K. Orientation-Resolving Pulsed Electron Dipolar High-Field EPR Spectroscopy on Disordered Solids: I. Structure of Spin-Correlated Radical Pairs in Bacterial Photosynthetic Reaction Centers. *J. Phys. Chem. B* **2007**, *111*, 6245−6262.
(28) Bogachev, A. V.; Kulik, L. V.; Bloch, D. A.; Bertsova, Y. V.; Fadeeva, M. S.; Verkhovsky, M. I. Redox Properties of the Prosthetic Groups of Na^+-Translocating NADH:Quinone Oxidoreductase. 1. Electron Paramagnetic Resonance Study of the Enzyme. *Biochemistry* **2009**, *48*, 6291−6298.
(29) Astashkin, A. V.; Elmore, B. O.; Fan, W.; Guillemette, J. G.; Feng, C. Pulsed EPR Determination of Distance between Heme Iron and FMN Centers in a Human Inducible Nitric Oxide Synthase. *J. Am. Chem. Soc.* **2010**, *132*, 12059−12067.
(30) Savitsky, A.; Dubinskii, A. A.; Zimmermann, H.; Lubitz, W.; Möbius, K. High-Field Dipolar Electron Paramagnetic Resonance (EPR) Spectroscopy of Nitroxide Biradicals for Determining Three-Dimensional Structures of Biomacromolecules in Disordered Solids. *J. Phys. Chem. B* **2011**, *115*, 11950−11963.
(31) Astashkin, A. V.; Rajapakshe, A.; Cornelison, M. J.; Johnson-Winters, K.; Enemark, J. H. Determination of the Distance between the Mo(V) and Fe(III) Heme Centers of Wild Type Human Sulfite Oxidase by Pulsed EPR Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 1942−1950.
(32) Savitsky, A.; Niklas, J.; Golbeck, J. H.; Möbius, K.; Lubitz, W. Orientation Resolving Dipolar High-Field EPR Spectroscopy on Disordered Solids: II. Structure of Spin-Correlated Radical Pairs in Photosystem I. *J. Phys. Chem. B* **2013**, *117*, 11184−11199.
(33) Razzaghi, S.; Qi, M.; Nalepa, A. I.; Godt, A.; Jeschke, G.; Savitsky, A.; Yulikov, M. RIDME Spectroscopy with Gd(III) Centers. *J. Phys. Chem. Lett.* **2014**, *5*, 3970−3975.
(34) Schweiger, A.; Jeschke, G. *Principles of Pulse Electron Paramagnetic Resonance*; Oxford University Press: Oxford, U.K., 2001.
(35) Hagelüken, G.; Ward, R.; Naimsith, J. H.; Schiemann, O. MtsslWizard: In silico Spin-Labelling and Generation of Distance Distributions in PyMOL. *Appl. Magn. Reson.* **2012**, *42*, 377−391.
(36) Hagelueken, G.; Abdullin, D.; Ward, R.; Schiemann, O. In Silico Spin Labelling, Trilateration and Distance-Constrained Rigid Body Docking in PyMOL. *Mol. Phys.* **2013**, *111*, 2757−2766.
(37) Stoll, S.; Lee, Y.-T.; Zhang, M.; Wilson, R. F.; Britt, R. D.; Goodin, D. B. Double Electron-Electron Resonance Shows Cytochrome P450cam Undergoes a Conformational Change in Solution upon Binding Substrate. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 12888−12893.
(38) Liu, H.; Naismith, J. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008**, 8−91.
(39) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006 − a comprehensive software package for analyzing pulsed ELDOR data. *Appl. Magn. Reson.* **2006**, *30*, 473−498.
(40) Masuya, F.; Tsubaki, M.; Makino, R.; Hori, H. EPR Studies on the Photoproducts of Ferric Cytochrome P450cam (CYP101) Nitrosyl Complexes: Effects of Camphor and Its Analogues on Ligand-Bound Structures. *J. Biochem.* **1994**, *116*, 1146−1152.
(41) Narr, E.; Godt, A.; Jeschke, G. Selective Measurements of a Nitroxide-Nitroxide Separation of 5 nm and a Nitroxide-Copper Separation of 2.5 nm in a Terpyridine-Based Copper(II) Complex by Pulse EPR Spectroscopy. *Angew. Chem., Int. Ed.* **2002**, *41*, 3907−3910.
(42) Bode, B. E.; Plackmeyer, J.; Prisner, T. F.; Schiemann, O. PELDOR Measurements on a Nitroxide-Labeled Cu(II) Porphyrin: Orientation Selection, Spin-Density Distribution, and Conformational Flexibility. *J. Phys. Chem. A* **2008**, *112*, 5064−5073.
(43) Lovett, J. E.; Bowen, A. M.; Timmel, C. R.; Jones, M. W.; Dilworth, J. R.; Caprotti, D.; Bell, S. G.; Wong, L. L.; Harmer, J. Structural Information from Orientationally Selective DEER Spectroscopy. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6840−6848.
(44) Lindgren, M.; Eaton, G. R.; Eaton, S. S.; Jonsson, B. H.; Hammarstrom, P.; Svensson, M.; Carlsson, U. Electron Spin Echo Decay as a Probe of Aminoxyl Environment in Spin-Labeled Mutants of Human Carbonic Anhydrase II. *J. Chem. Soc., Perkin Trans.* **1997**, *2*, 2549−2554.
(45) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-band Frequencies. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10762−10773.
(46) Spindler, P. E.; Glaser, S. J.; Skinner, T. E.; Prisner, T. F. Broadband Inversion PELDOR Spectroscopy with Partially Adiabatic Shaped Pulses. *Angew. Chem., Int. Ed.* **2013**, *52*, 3425−3429.

DOI: 10.1021/acs.jpcb.5b02118
J. Phys. Chem. B 2015, 119, 13534−13542

Supporting Information

Comparison of PELDOR and RIDME for Distance Measurements between Nitroxides and Low-Spin Fe(III) Ions

*Dinar Abdullin, Fraser Duthie, Andreas Meyer, Elisa S. Müller, Gregor Hagelueken, and Olav Schiemann**

Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany

Table of Contents

1. Biochemical characterization of the cytochrome P450cam mutant $C58R_1$

Figure S1a shows an SDS-PAGE analysis of cytochrome P450cam samples before (load) and after gelfiltration (elution). The samples eluted in a single symmetric peak on a Superdex 200 16/60 gelfiltration column (Figure S1b). The observed elution volume of ~87 ml is in agreement with a monomeric cytochrome P450cam. Circular dichroism (CD) spectra of the wild type cytochrome P450cam and its mutant $C58R_1$ recorded by means of the Jasco J-810 spectrometer are shown in Figure S1c. Both spectra show no significant difference with respect to each other and to spectra observed in previous works.[1] This indicates that the mutation introduced in $C58R_1$ does not change the structure of the protein.

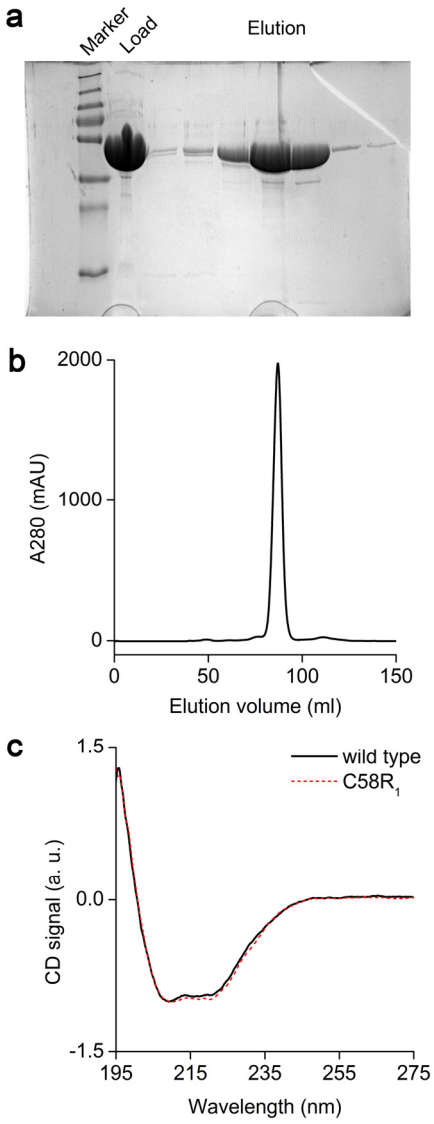


Figure S1. a) SDS-PAGE analysis of purified cytochrome P450cam mutant C58. **b)** Size-exclusion chromatogram of cytochrome P450cam mutant C58. **c)** CD spectra of the wild type cytochrome P450cam and its mutant $C58R_1$.

2. CW EPR measurements on the cytochrome P450cam mutant $C58R_1$

Continuous wave EPR measurements were carried out using a Bruker ELEXSYS E580 EPR spectrometer. The EPR spectra of the $C58R_1$ samples mutants were recorded at 293 K and 80 K with a sampling time of 20.48 ms, a microwave power of 20 mW, a modulation amplitude of 0.1 mT, and a modulation frequency of 100 kHz.

CW EPR spectra of the $C58R_1$ samples are shown in Figure S2. The spectra recorded at 293K feature the characteristic spectra of protein-bound nitroxides. The labelling degree was estimated to be >80% comparing the integral intensities of the nitroxide spectra corresponding to the $C58R_1$ samples and to a reference MTSSL solution. A small fraction (<5%) of remaining unbound spin labels is visible in the spectrum of $C58R_1$-H_2O as sharp peaks in the low- and high-field regions. The obtained nitroxide spectra are typical for significantly immobilized MTSSL.[2,3] Moreover, two distinct components are observed for the low- and high-field regions of the nitroxide spectrum in $C58R_1$-D_2O (Figure S2a, stars) suggesting the existence of two different populations for the spin label. Note that very similar nitroxide spectra were observed by Lopez *et al.* for the T4 lysozyme W138A mutants.[2] In that work, the spectra were simulated assuming an anisotropic motion of MTSSL with an effective rotational correlation time of ~1.5 ns and using two different populations of MTSSL.

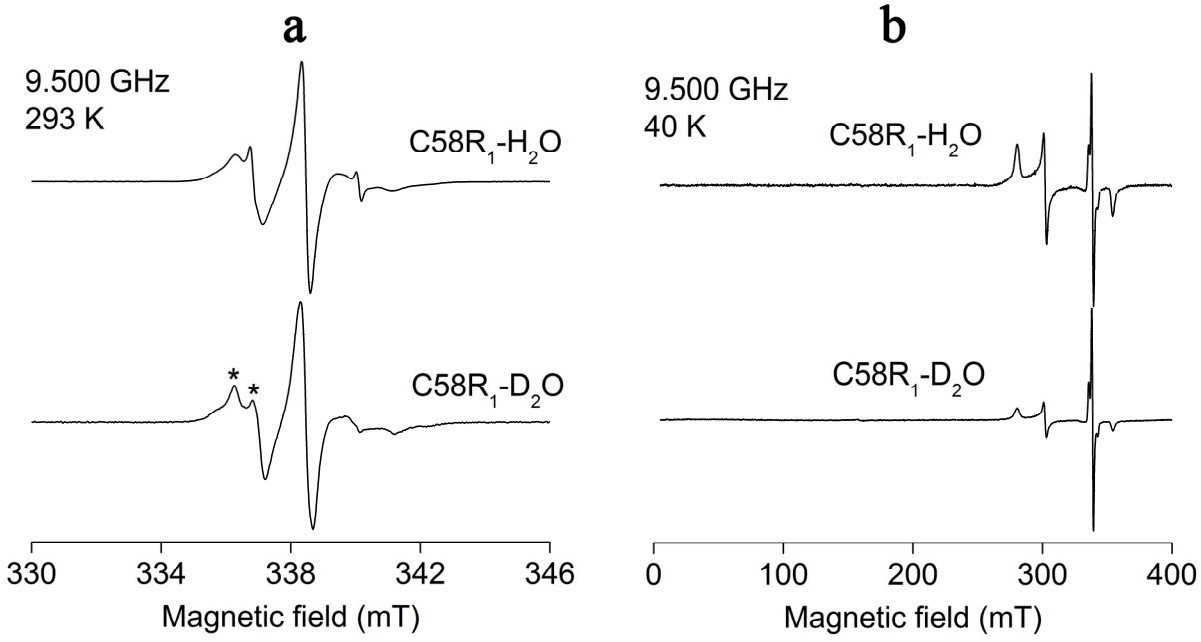


Figure S2. CW spectra of the $C58R_1$-H_2O and $C58R_1$-D_2O samples recorded at **a)** 293 K and **b)** 40 K. The components of the nitroxide spectrum of $C58R_1$-D_2O corresponding to two different MTSSL populations are marked by stars.

The spectra recorded at 40 K consist of two contributions: an intense signal at $g \sim 2.00$ which corresponds to the nitroxide and a broader low-amplitude signal which is described by a *g*-

tensor (g_1, g_2, g_3) = (2.42, 2.25, 1.91) originating from the low-spin Fe^{3+} ion. One can see that the intensity of the Fe^{3+} signal varies for the two different samples, which is attributed to different contents of the heme group in $C58R_1$-H_2O and $C58R_1$-D_2O samples.

3. UV-Vis measurements on the cytochrome P450cam mutant $C58R_1$

UV-Vis measurements were performed by means of a NanoDrop 2000 spectrometer (Thermo Scientific). The spectra of the $C58R_1$-H_2O and $C58R_1$-D_2O samples are illustrated in Figure S3. Both spectra contain a protein peak at 280 nm and the Soret band at 417 nm corresponding to the ferric heme cofactor. The ratio of the amplitudes of these two peaks (A417/A280) determines the content of heme cofactor per cytochrome molecule. The A417/A280 ratios of 1.3 and 1.0 were found for $C58R_1$-H_2O and $C58R_1$-D_2O, respectively. Based on these values one can conclude that the $C58R_1$-H_2O sample has a higher content of heme cofactor than the $C58R_1$-D_2O. Although we did not find in the literature the A417/A280 ratio corresponding to a 100% content of the heme cofactor in CYP101, it was estimated by means of:

$$\frac{A_{417}}{A_{280}} = \frac{\varepsilon_{417}}{\varepsilon_{280}} \qquad \text{(S1)}$$

where ε_{417} and ε_{280} are the extinction coefficients of CYP101 at 414 nm and 280 nm, respectively. Using the literature values of ε_{417} = 102 $mM^{-1}cm^{-1}$ and ε_{280} = 40.91 $mM^{-1}cm^{-1}$,[4] the ideal A417/A280 ratio is estimated to be 2.5. Note, this value is higher than the experimental A417/A280 ratios of 1.0 – 1.6 reported in previous works.[5-7] Thus, the amount of bound heme cofactor in the samples can be estimated to be 52% for $C58R_1$-H_2O 40% for $C58R_1$-D_2O.

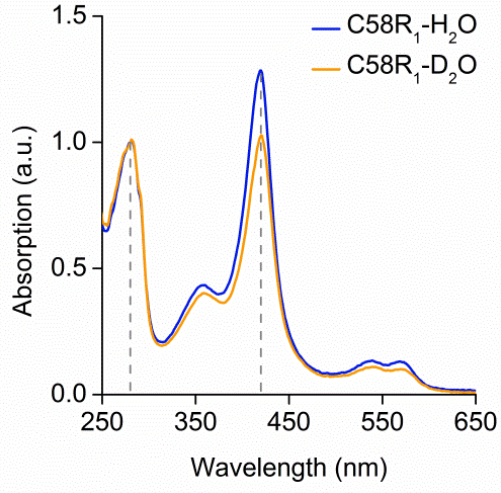


Figure S3. UV-Vis spectra of the $C58R_1$-H_2O and $C58R_1$-D_2O samples. The wavelengths of 280 nm and 417 nm are shown by dashed lines.

4. Pulsed EPR measurements on the cytochrome P450cam mutant $C58R_1$

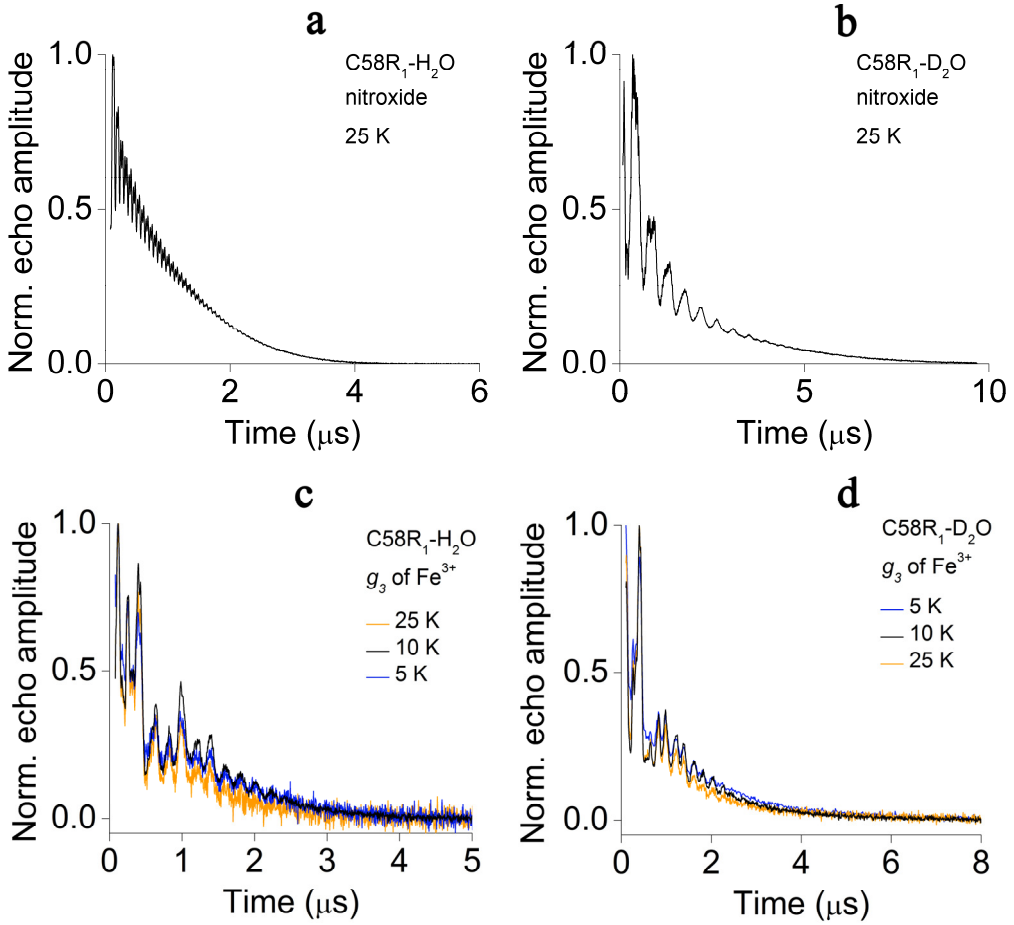


Figure S4. 2-pulse ESEEM of **a)** the nitroxide of $C58R_1$-H_2O at 25 K, **b)** the nitroxide of $C58R_1$-D_2O at 25 K, **c)** the Fe^{3+} g_3-component of $C58R_1$-H_2O at 5, 10 and 25 K, **d)** the Fe^{3+} g_3-component of $C58R_1$-D_2O at 5, 10 and 25 K.

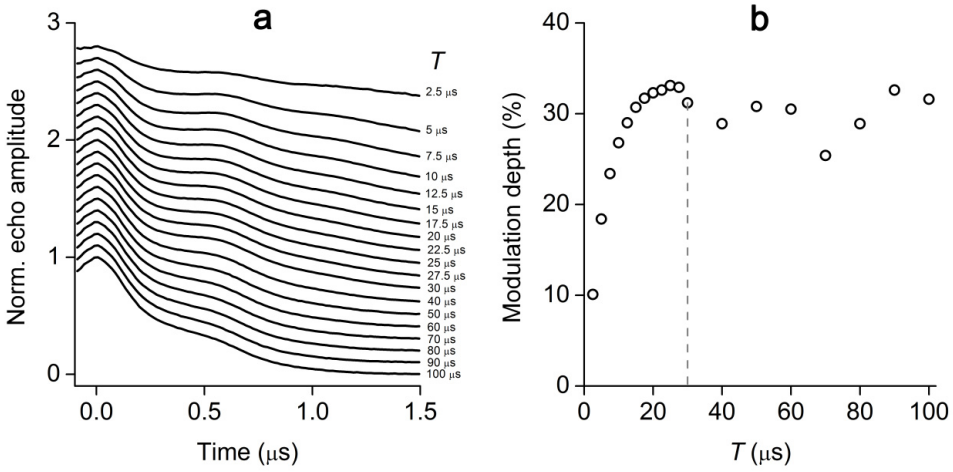


Figure S5. a) Q-band RIDME time traces recorded on the $C58R_1$-H_2O sample with the different inter-pulse intervals *T*. **b)** Dependence of the modulation depth parameter on *T*.

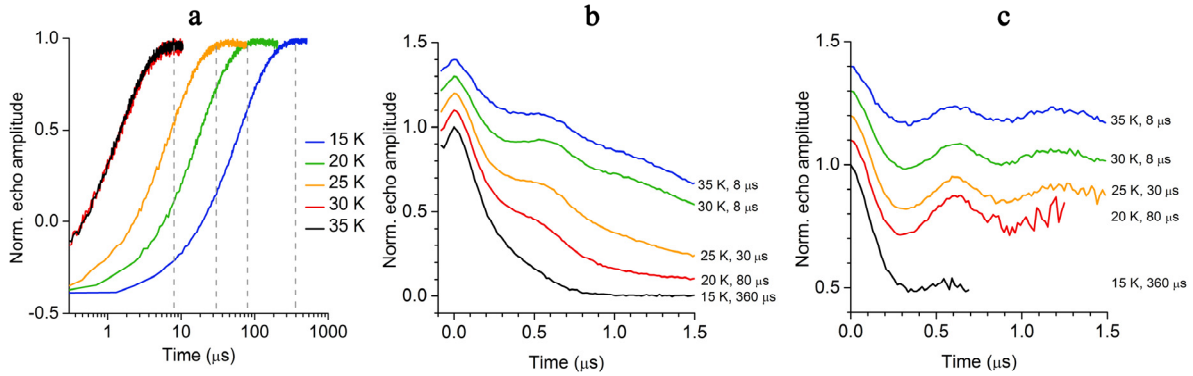


Figure S6. a) Q-band inversion recovery for the Fe^{3+} spin center in C58R$_1$-H$_2$O recorded at different temperatures. The dashed lines correspond to the *T* values chosen for the RIDME experiments. **b)** The Q-band RIDME time traces recorded on the C58R$_1$-H$_2$O sample at different temperatures and **c)** the corresponding background-subtracted time traces. A good compromise between the modulation depth and the signal's length is achieved at temperature of 25 K. At temperatures >25 K the optimal *T* value is only a few microseconds leading to violation of the requirement $T >> t$ and consequently to the reduction of the modulation depth. At temperatures <25 K the RIDME signal rapidly decays to zero due to relaxation of nitroxide spins during *T*.

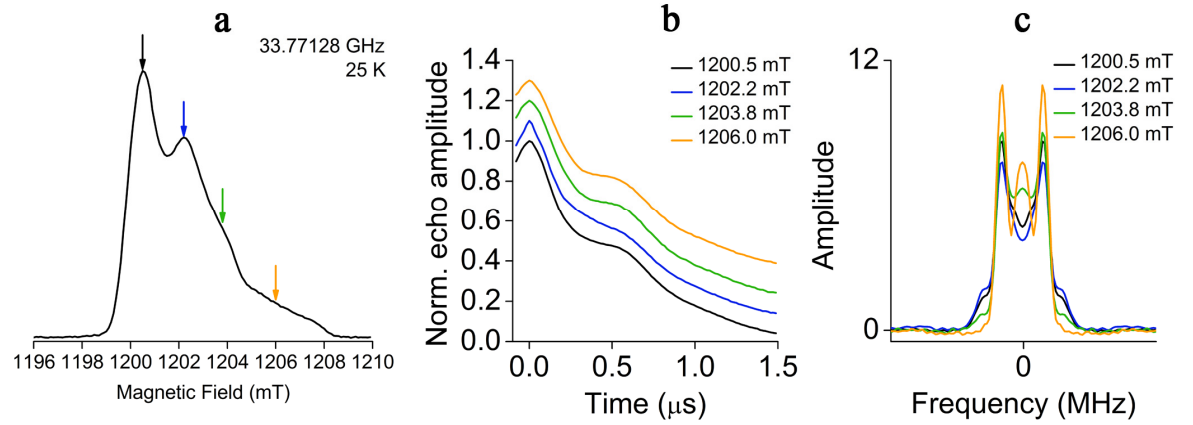


Figure S7. a) Detection positions on the nitroxide spectrum, **b)** the corresponding RIDME time traces, and **c)** their FFTs.

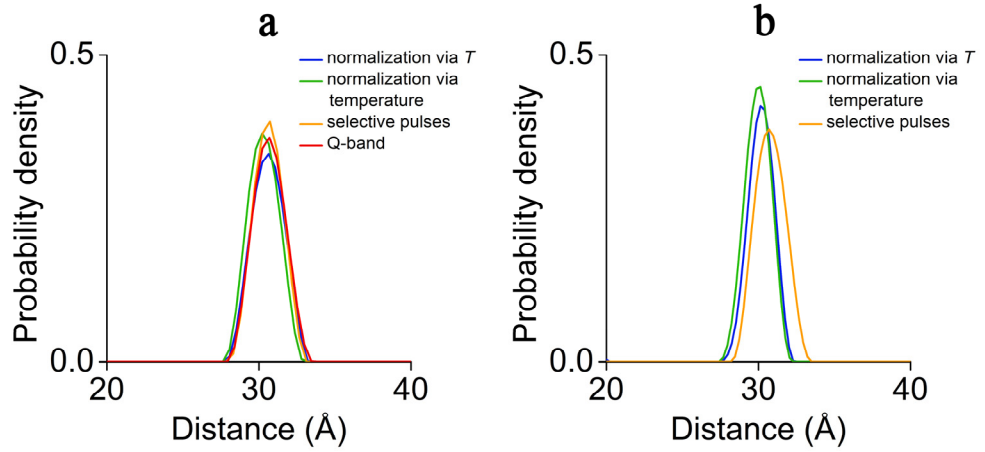


Figure S8. Overlay of the RIDME-derived distance distributions for different procedures of ESEEM suppression. The overlay is shown for **a)** the C58R$_1$-H$_2$O and **b)** the C58R$_1$-D$_2$O samples.

References

(1) Manna, S. K.; Mazumdar S. Role of Threonine 101 on the Stability of the Heme Active Site of Cytochrome P450cam: Multiwavelength Circular Dichroism Studies. *Biochemistry* **2006,** *45,* 12715-12722.

(2) Lopez, C. J.; Yang, Z.; Altenbach, C.; Hubbell, W. L. Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4306–E4315.

(3) Beier, C.; Steinhoff, H.-J. A Structure-Based Simulation Approach for Electron Paramagnetic Resonance Spectra Using Molecular and Stochastic Dynamics Simulations. *Biophys. J.* **2006**, *91*, 2647-2664.

(4) Gunsalus, I. C.; Wagner, G. C. Bacterial P-450cam methylene monooxygenase components: cytochrome m, putidaredoxin, and putidaredoxin reductase. *Methods Enzymol.* **1978**, *52*, 166-188.

(5) Contzen, J.; Jung, C. Step-Scan Time-Resolved FTIR Spectroscopy of Cytochrome P-450$_{cam}$ Carbon Monoxide Complex: A Salt Link Involved in the Ligand-Rebinding Process. *Biochemistry* **1998**, *37*, 4317-4324.

(6) Sakurai, K.; Shimada, H.; Hayashi, T.; Tsukihara, T. Substrate binding induces structural changes in cytochrome P450cam. *Acta Cryst.* **2009**, *F65*, 80-83.

(7) Lee, Y.-T. ; Glazer, E. C.; Wilson, R. F.; Stout, C. D.; Goodin, D. B. Three Clusters of Conformational States in P450cam Reveal a Multistep Pathway for Closing of the Substrate Access Channel. *Biochemistry* **2011**, *50*, 693-703.

[P7] Post-synthetic spin labeling of RNA via click chemistry for PELDOR measurements

Reprinted with permission from

Mark Kerzhner, Dinar Abdullin, Jennifer Więcek, Hideto Matsuoka, Gregor Hagelueken, Olav Schiemann, Michael Famulok, "Post-synthetic Spin Labeling of RNA via Click Chemistry for PELDOR Measurements", *Chem. Eur. J.* **2016**, *22*, 12113-12121.
DOI: 10.1002/chem.201601897

Own contribution to the manuscript:

- Performing CD, CW-EPR, PELDOR and RIDME measurements
- Analysis of the PELDOR and RIDME data, comparing these methods
- Writing the manuscript

ChemPubSoc Europe

DOI: 10.1002/chem.201601897

CHEMISTRY A European Journal Full Paper

Spin Labeling |*Hot Paper*|

Post-synthetic Spin-Labeling of RNA through Click Chemistry for PELDOR Measurements**

Mark Kerzhner[+],[a] Dinar Abdullin[+],[b] Jennifer Więcek,[b] Hideto Matsuoka,[b] Gregor Hagelueken,[b] Olav Schiemann,*[b] and Michael Famulok*[a, c]

Abstract: Site-directed spin labeling of RNA based on click chemistry is used in combination with pulsed electron-electron double resonance (PELDOR) to benchmark a nitroxide spin label, called here **dU̧**. We compare this approach with another established method that employs the rigid spin label **Çm** for RNA labeling. By using CD spectroscopy, thermal denaturation measurements, CW-EPR as well as PELDOR we analyzed and compared the influence of **dU̧** and **Çm** on a self-complementary RNA duplex. Our results demonstrate that the conformational diversity of **dU̧** is significantly reduced near the freezing temperature of a phosphate buffer, resulting in strongly orientation-selective PELDOR time traces of the **dU̧**-labeled RNA duplex.

Introduction

As a carrier of genetic information, RNA plays a crucial role in living organisms. Its ability to form a variety of three-dimensional structures enables RNA to perform various functions. Riboswitches and small RNAs, such as siRNA and miRNA, control and regulate gene expression, ribozymes catalyze the maturation of mRNA, and rRNA catalyzes the polypeptide bond formation in ribosomes. Many of these functions occur in a highly dynamic fashion, involving structural rearrangements of RNA. To investigate the structure and conformational changes of RNA, experimental methods are required. One of the methods that holds potential in this field is pulsed electron-electron double resonance (PELDOR, also known as double electron-electron resonance or DEER) spectroscopy.[1] This technique allows the determination of structural constraints for biomolecules by measuring the distances between two or more spin centers in the biomolecule on the 1.5–15 nm length scale.[2] Given that many biomolecules are diamagnetic, the spin centers are usually introduced through site-directed spin labeling.[3] Commonly, nitroxides are used as the spin labels because of their stability and well-known EPR properties. The introduction of nitroxide spin labels into the oligonucleotide structure can be performed either during the oligonucleotide synthesis[4] or through post-synthetic labeling of pre-functionalized sites of the oligonucleotide.[5] Post-synthetic spin labeling has the important advantage that it avoids subjecting the nitroxide to potentially reducing conditions during the oligonucleotide synthesis. The development of post-synthetic methods is therefore of high interest, as revealed by a number of recently developed modifications that have been applied to all three structural units of oligonucleotides: the sugar moiety,[6] the phosphate backbone,[7] and the nucleobases.[8] In case of RNA, such modifications include the 2′-amino-modified pyrimidine nucleotides,[6b] the thiolate-modified phosphate backbone,[8b] the 4-thiouridine nucleotide,[9] and the exocyclic amino group of nucleobases.[7b] An alternative approach for labeling RNA nucleobases that can be used for long natural RNA strands, employs the site-selective attachment of a linker containing an aliphatic amino group to the target nucleotide and the subsequent coupling of a spin label to this linker.[10] This variety of post-synthetic spin labeling methods allows for the incorporation of spin labels with different structural properties. One important property of spin labels is the flexibility of their molecular linker connecting the nitroxide with the RNA. On one hand, flexible linkers usually result in relatively broad distance distributions between two labels, often making the interpretation of PELDOR-derived distances in terms of the RNA structure less precise. However, they allow for a relatively straightforward PELDOR data analysis because of the absence of orientation selectivity effects. On the other hand, rigid linkers yield narrower inter-label distance distributions, but require a more demanding analysis of the orientation-selective

[a] *M. Kerzhner,[+] Prof. Dr. M. Famulok*
Life & Medical Sciences Institute
Chemical Biology & Medicinal Chemistry Unit
c/o Kekulé-Institut für Organische Chemie und Biochemie
University of Bonn, Gerhard-Domagk-Str. 1, 53121 Bonn (Germany)
E-mail: m.famulok@uni-bonn.de

[b] *D. Abdullin,[+] J. Więcek, Dr. H. Matsuoka, Dr. G. Hagelueken, Prof. Dr. O. Schiemann*
Institute for Physical and Theoretical Chemistry, University of Bonn
Wegelerstr. 12, 53115 Bonn (Germany)
E-mail: schiemann@pc.uni-bonn.de

[c] *Prof. Dr. M. Famulok*
Max-Planck Fellow Group Chemical Biology
Center of Advanced European Studies and Research
Ludwig-Erhard-Allee 2, 53175, Bonn (Germany)

[+] *These authors contributed equally to this work.*

[**] *PELDOR = pulsed electron-electron double resonance.*

Supporting information for this article can be found under http://dx.doi.org/10.1002/chem.201601897.

PELDOR data. A typical example of a flexible spin label is the 3-(2-iodoacetamido)proxyl spin label.[5b] By using this label, the structure of a 70 kDa RNA–protein complex was solved by means of NMR and PELDOR-derived long-range distance constraints.[10] In contrast, one of the most rigid spin labels, introduced into both DNA and RNA, is the spin label **Ç** ("C-spin")[11] and its methylated analogue **Çm**.[12] Both labels were successfully employed for PELDOR measurements, providing not only the distance between the labels but also their relative orientation.[13] This information has enabled the investigation of the conformational flexibility of DNA, revealing a huge potential of rigid spin labels in elucidation of the oligonucleotides' dynamics.[14] Unfortunately, the synthesis of **Çm** is complex and time consuming and the incorporation of **Çm** into RNA occurs during solid-phase synthesis. Therefore, it is of interest to develop alternative post-synthetic spin-labeling approaches that yield spin labels of similar rigidity to **Çm** that can be incorporated into RNA in a versatile and efficient way.

A frequently used method for oligonucleotide functionalization is click chemistry, one form of which involves the reaction between an alkyne and an azide that proceeds by Cu^{I}-catalyzed Huisgen–Meldal–Sharpless [3+2] cycloaddition.[15] Click chemistry has been used previously to attach spin labels to DNA. For example, Seela et al. reported the spin labeling of alkyne-modified DNA in solution, in which 4-azido-2,2,6,6-tetramethylpiperidine1-oxyl (4-azido-TEMPO) reacted with either 7-deaza-7-ethynyl-2′-deoxyadenosine or 5-ethynyl-2′-deoxyuridine.[16] Others reported the spin labeling of DNA on solid support by using click chemistry in which the nitroxide-functionalized azide **1** (Figure 1) was efficiently incorporated into DNA containing 5-ethynyl-2′-dU.[17] EPR measurements on this spin-labeled DNA at room temperature revealed that the label is flexible to some degree.[17] However, ab initio calculations have shown that the triazole–pyrimidine bond has a preferred conformation[18] and thus one could expect a reduced mobility of the incorporated spin label.

Here, we applied click chemistry for a highly efficient post-synthetic labeling of RNA. The azide-functionalized nitroxide **1** was introduced into the alkyne-modified self-complementary RNA sequence **2** to yield the spin-labeled nucleobase denoted here as **dU̧** (Figure 1). The new labeling procedure was benchmarked in relation to the previously reported procedure for **Çm**.[12] For this purpose, we labeled the similar RNA sequence **3** with **Çm** (Figure 1). We compare both RNA sequences by thermal denaturation, CD and EPR spectroscopic analyses. In addition, Q-band PELDOR measurements were conducted on both RNA duplexes to demonstrate the advantages and disadvantages of the spin label **dU̧** with respect to distance measurements.

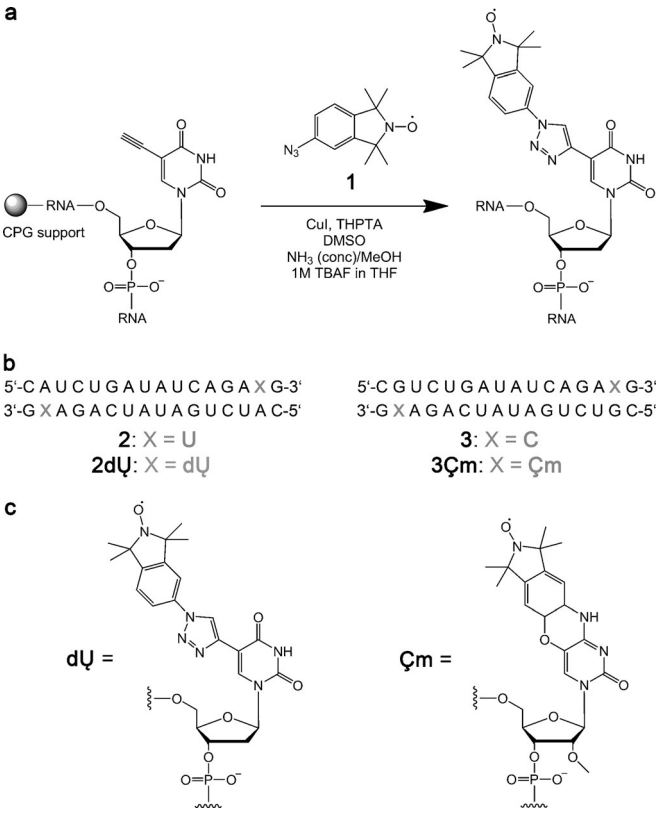


Figure 1. a) Spin labeling of RNA with the azide-functionalized nitroxide on CPG support by using click chemistry. b) Non-modified and spin-labeled self-complementary RNA sequences. The labeling positions are shown in gray. c) The structures of the incorporated spin-labeled nucleotides.

Results and Discussion

Synthesis and spin labeling of RNA

5-Ethynyl-2′-dU modified RNA **2** covalently attached to the CPG support was incubated in a solution of **1**, CuI, and the Cu^{I} stabilizing ligand tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) in dimethyl sulfoxide (DMSO), followed by deprotection and cleavage of the RNA strand from the solid support (Figure 1). Note that copper(I)-catalyzed azide–alkyne [3+2] cycloadditions usually generate the catalytically active Cu^{I} species through reduction of Cu^{II} by ascorbate in situ.[19] Thus, due to the reducible nature of **1**, CuI was used directly. The THPTA ligand was used to prevent the oxidation or disproportionation of Cu^{I}.[20] An important requirement of the click reaction was a thorough degassing of the solvent. Performing the reaction in non-degassed solvents led to very low conversion in the click reaction, as revealed by LC-MS analysis. This observation clearly demonstrates a strong sensitivity of Cu^{I} towards oxygen, even in the presence of THPTA. The LC-MS analysis of the crude reaction mixture after cleavage and deprotection of the RNA displayed a nearly quantitative conversion of the click reaction (see the Supporting Information). No reduced nitroxide species was observed, indicating that the spin label was stable towards the RNA deprotection conditions.

Çm was synthesized and incorporated into the RNA strand **3** by using the procedure described previously.[12]

The synthesis of the **dU̧** precursor **1** was significantly easier than the synthesis of **Çm**, because it can be completed in 7 instead of 16 steps and requires considerably less starting material to yield a similar amount of the final product **1** as compared with **Çm**. The final yield of spin labeled and purified RNA was 18 nmol (9%) and 23 nmol (12%) for **Çm** and **1**, re-

spectively, for a 0.2 μmol scale RNA synthesis. According to CW-EPR analysis, both **2dU̇** and **3Çm** RNA strands were labeled quantitatively (see the Supporting Information).

CD measurements

CD spectra of the non-modified (**2** and **3**) and spin-labeled (**2dU̇** and **3Çm**) RNA duplexes revealed that all duplexes adopt an A-form helix (Figure 2). All spectra display the characteristic absorption profile with a minimum at 210 nm and a maximum at 260 nm.[21] The spectrum of **2dU̇** displays only a 5 nm shift of the minimum to shorter wavelengths and a slight increase of both extrema as compared with **2**. This reveals that **dU̇** has a minor effect on the secondary structure of the RNA strand. In contrast, the introduction of **Çm** in **3** leads to more pronounced changes in the CD spectrum: Both extrema of the spectrum are shifted by 2 nm to longer wavelengths and their amplitudes are reduced. These effects were also observed previously for **Çm**-labeled RNA.[12] Thus, the introduction of **Çm** into **3** may cause some local perturbations in the helix structure.

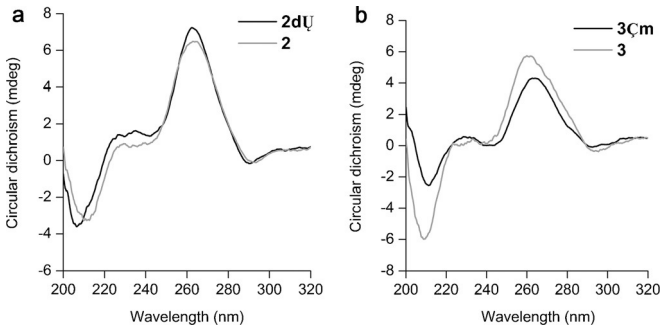


Figure 2. CD spectra of a) **2** (gray) and **2dU̇** (black), b) **3** (gray) and **3Çm** (black).

Measurements of thermal denaturation

The influence of the spin labels on the stability of the RNA duplexes was investigated by UV melting experiments (Figure S3). A decrease in the melting temperature (T_m) from 66 °C for **2** to 58 °C for **2dU̇** was observed. Taking into account that the RNA duplex contains two spin labels, the decrease in T_m caused by each label amounts to approximately −4 °C. A similar decrease in T_m was previously reported for **dU̇** in DNA.[17] Only a 1 °C difference in T_m was found between **3** (T_m = 73 °C) and **3Çm** (T_m = 72 °C). As was shown previously, the 2′-OMe group of **Çm** has a positive effect on the overall stability of the **Çm**-modified RNA, probably because the ribose is locked in the C3′-*endo* conformation.[22] Thus, the destabilizing effect of the spin-label modification is counterbalanced by the 2′-ribose methoxy modification.

CW-EPR measurements

Incorporation of the nitroxide spin labels **dU̇** and **Çm** into the respective RNA duplexes **2dU̇** and **3Çm** was verified by room-temperature CW-EPR spectroscopic analysis, wherein the spectra showed line patterns typical of a partially immobilized nitroxide (Figure 3). Interestingly, the spectra of both RNA duplexes display a superposition of two contributions corresponding to nitroxide centers of different mobilities. To characterize these centers, the experimental spectra were simulated by using the program MultiComponent[23] (Figure 3 and Table S1). Effective rotational correlation times τ_c of 1.8 and 0.9 ns were determined for the nitroxide spins in **2dU̇**, and of 7.9 and 0.8 ns for **3Çm**. For both duplexes, the nitroxide fraction with longer τ_c and therefore smaller mobility dominates both spectra (93 %). The remaining 7 % of the spin labels correspond to more mobile nitroxide centers with shorter τ_c. For simplicity, these two fractions of the spin labels are called here "slow-motion" and "fast-motion" fractions. The τ_c values of the slow-motion fraction are typical for the corresponding spin labels attached to double helices.[12,17] Therefore, this fraction is attributed to the RNA-bound spin labels. The origin of the fast-motion fraction is less clear. Given that the RNA single strands are self-complementary and thus the sites of the label are identical on both ends of each duplex, the appearance of the second fraction cannot be due to two different sites of labeling. The presence of single strands could be a possible reason, but the relatively high melting temperatures of **2dU̇** and **3Çm** make this unlikely. Furthermore, the presence of free labels in

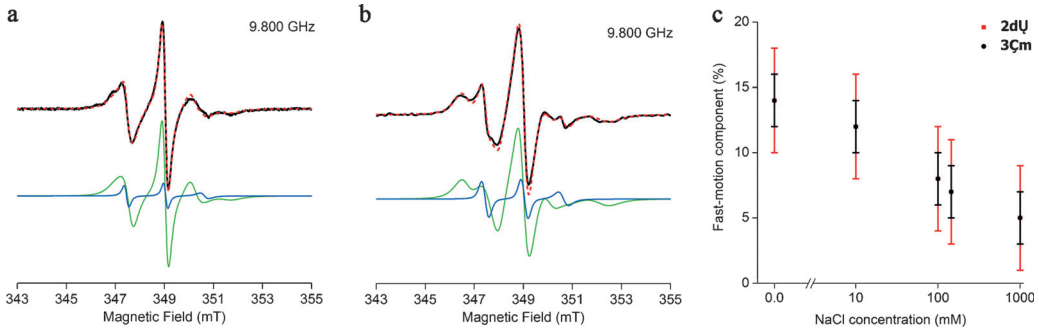


Figure 3. Experimental (black lines) and simulated (red dashes) CW-EPR spectra of a) **2dU̇** and b) **3Çm**. The concentration of NaCl in both samples was 145 mM. The simulated spectra of the slow-motion and fast-motion fractions of the spin labels are shown by green and blue lines, respectively. c) The content of the fast-motion fraction of the spin labels in **2dU̇** and **3Çm** as a function of the NaCl concentration.

the samples as the origin for the mobile spin labels is not likely, considering that different synthesis strategies were used. Moreover, the rotational correlation times of the fast-motion fraction are too large for a free label (0.8–0.9 ns here as compared to ca. 0.1 ns for the free labels). Notably, an increase of the NaCl concentration from 0 to 1 M led to a gradual decrease of the fast-motion fraction from 14 to 5% in both samples (Figure 3c and Figure S4). Thus, a possible explanation for the fast-motion fraction might be fraying of the RNA duplex ends, which is known to be reduced upon increasing salt concentrations.[24] In this case, the fast-motion fraction of the spin labels can be assigned to RNA duplexes, in which the ends are unpaired and move more freely in solution.

The rotational correlation times of the main nitroxide fraction in **2dU** ($\tau_c = 1.8$ ns) and **3Çm** ($\tau_c = 7.9$ ns) reveal that **dU** has a higher mobility than **Çm**. Note that the mobility of both spin labels consists of the motion of the entire RNA, segmental motion of the backbone, and the motion of the spin labels' linkers. Given that the RNA sequences **2dU** and **3Çm** are almost identical, the difference in the mobility of **dU** and **Çm** should originate mostly from the different mobility of their linkers.

PELDOR measurements and analysis

The spin labeling of RNA with **dU** and **Çm** was conducted to enable both CW-EPR and PELDOR measurements. The potential of such measurements was explored by performing Q-band PELDOR experiments on **2dU** and **3Çm**. Previously, X- and G-band PELDOR studies on **Ç** and **Çm** were conducted, revealing a strong dependence of the PELDOR time traces on the relative orientation of these rigid labels.[13,14] Here, this is also observed for the time traces of **3Çm** (Figure 4a). More unexpectedly, the PELDOR time traces of **2dU** were also found to be strongly orientation selective (Figure 4b). This is reflected in a varying modulation depth λ between 0.1 and 0.2 and varying modulation frequencies (Figure S5) in dependence of the pump and probe positions.

The analysis of the PELDOR time traces of **2dU** and **3Çm** was performed by using the program PeldorFit.[25] This program simulates the PELDOR time traces for a simplified geometric model of the nitroxide spin pairs (Figure 5) and subsequently calculates the root-mean-square deviation (RMSD) between the simulated and experimental time traces. The model is optimized by means of a genetic algorithm until the corresponding RMSD reaches its minima. The parameters of the model that are optimized are three spherical coordinates (r, ξ, φ), which describe the distance vector relative to one of the spins, and three Euler angles (α, β, γ), which describe the relative position and orientation of the two magnetic frames, respectively. To account for the flexibility of the RNA molecules carrying the spins, all six parameters are assumed to follow a normal distribution. Thus, each of the parameters is described by a mean value (r, ξ, φ, α, β, γ) and a standard deviation (σ_r, σ_ξ, σ_φ, σ_α, σ_β, σ_γ).

Applying PeldorFit to the present PELDOR data, good reproduction of the experimental time traces by the simulated ones

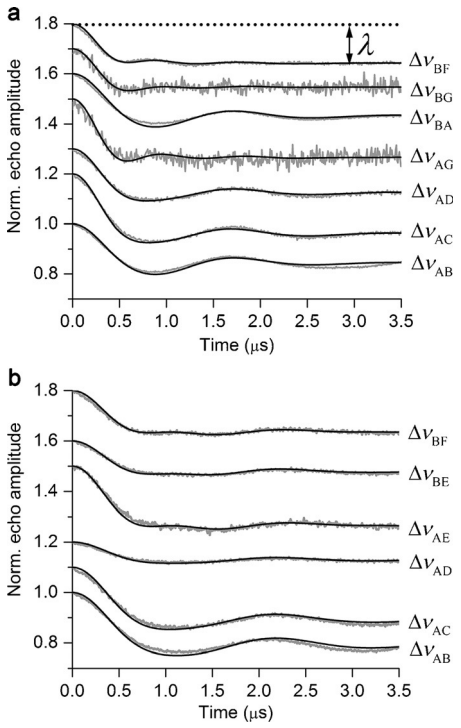


Figure 4. Q-band PELDOR time traces (gray) acquired at different frequency offsets are overlaid with their simulations (black) obtained by means of the program PeldorFit[27] for a) **3Çm** and b) **2dU**. The notations of the different frequency offsets are given in the Experimental Section.

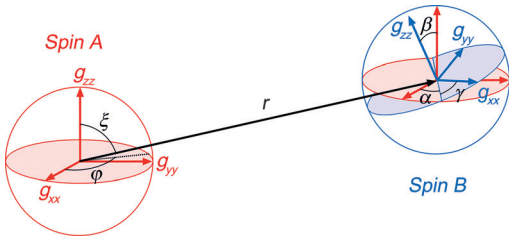


Figure 5. The model of a spin pair used in PeldorFit.

was achieved (Figure 4). The parameters of the corresponding optimized models are listed in Table 1. A mean distance of 4.68 nm and a standard deviation of 0.24 nm characterize the inter-label distance distribution for **2dU**. The same parameters for **3Çm** are 4.30 and 0.24 nm, respectively. To estimate, how

Table 1. Optimized geometric parameters of the PeldorFit model for **2dU** and **3Çm**.

Parameters	**2dU**[a]	**3Çm**[a]
r, σ_r (nm)	4.68 (0.1), 0.24 (0.1)	4.30 (0.1), 0.24 (0.1)
ξ, σ_ξ (°)	38 (10), 12 (10)	34 (10), 3 (10)
φ, σ_φ (°)	8 (40), 1 (50)	17 (40), 29 (60)
α, σ_α (°)	19 (40), 2 (35)	159 (30), 46 (60)
β, σ_β (°)	19 (25), 18 (15)	44 (25), 15 (30)
γ, σ_γ (°)	66 (50), 1 (30)	174 (40), 34 (60)

[a] The confidence interval of each parameter is given in parentheses.

defined the obtained distances and other optimized parameters of the model are, dependences of RMSD on the mean value and standard deviation of each parameter was calculated (Figure 6). While acquiring the RMSD dependence for a certain parameter, all other parameters were set to their optimized values. The latter condition allowed us to avoid recording the RMSD plot for the entire 12-dimentional parameter space, which would be very time expensive. Parameter ranges, in which 110% of the minimal RMSD are reached, were used to determine approximate confidence intervals for the optimized parameters (Figure S6 and Table 1). The RMSD plots calculated for r and σ_r revealed a 0.1 nm precision for the determined distances. Analogous RMSD plots of ξ and β display moderate uncertainties of approximately 10° and approximately 25°, respectively. In contrast, the angles φ, α, and γ are less well defined: The RMSD plots display only shallow maxima for α and γ and almost no dependence on φ. This result can be attributed to the relatively small difference between the *xx* and *yy* components of the **g**- and **A**-tensors at Q-band frequencies. Another reason for the poorly defined angles could be a twist-stretch-like motion of the RNA duplexes, which was observed previously for DNA duplexes.[14] In addition, the angles α and γ are slightly better defined for **3 Çm** as compared with **2 dU̧**, which reveals a small difference in the conformational distribution of the corresponding spin labels. Note that the RMSD plots of φ, α, and γ display multiple minima, which is due to the presence of several symmetry-related sets of these angles that are indistinguishable in PELDOR experiments. The reason for the symmetric solutions is invariance of the **g**- and **A**-tensors towards inversion of their axes. This symmetry results in 16 possible combinations of ξ, φ, α, β, and γ (Tables S3 and S4). However, given that only two angular parameters are varied in each RMSD plot and only half of the entire angular ranges are considered, one cannot see all 16 symmetric solutions in Figure 6.

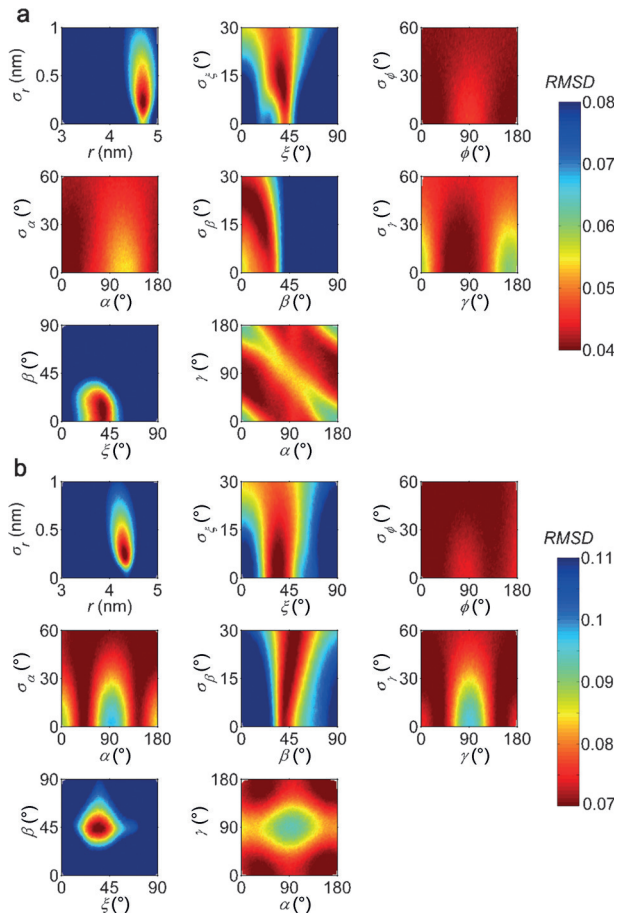


Figure 6. RMSD between experimental and simulated PELDOR time traces as a function of the geometric parameters of PeldorFit model for a) **2 dU̧** and b) **3 Çm**.

PELDOR data interpretation

The CW-EPR spectra of **2 dU̧** and **3 Çm** revealed, as expected, that the nitroxide group of **dU̧** is more flexible than the nitroxide group of **Çm** at room temperature (Figure 3). However, the PELDOR time traces of **2 dU̧** display only slightly less orientation selectivity than the time traces of **3 Çm** and the PELDOR-derived distance distributions between the two **dU̧** labels in **2 dU̧** and the two **Çm** labels in **3 Çm** have the same widths for the frozen samples (Table 1). Assuming that the overall mobility of both RNA molecules is very similar, the latter observation supports a similar conformational heterogeneity of **Çm** and **dU̧** at the freezing temperature of the solvent. This, in turn, shows that lowering the temperature from room temperature to the freezing point of the solvent may significantly reduce the conformational diversity of **dU̧**.

To study the dynamics of **dU̧** in more detail, we performed DFT calculations of its spin-labeled nucleobase (Figure 7). The calculations yielded an almost planar geometry for the spin-labeled nucleobase. By using this optimized structure, a relaxed scan of the energy profiles for the dihedral angles χ_1 (C_1-C_2-C_3-

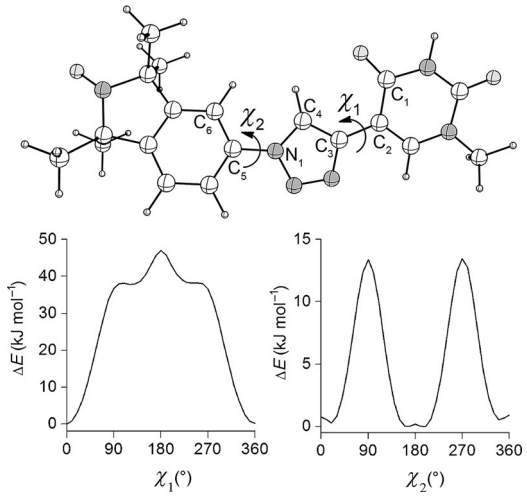


Figure 7. DFT structure of the **dU̧** nucleobase. The carbon, oxygen, nitrogen, and hydrogen atoms are drawn as white, gray, light gray, and small white spheres, respectively. The dihedral angles χ_1 (C_1-C_2-C_3-C_4) and χ_2 (C_4-N_1-C_5-C_6) are shown by arrows. Torsional energy profiles for these angles are shown below.

C_4) and χ_2 (C_4-N_1-C_5-C_6) was carried out (Figure 7). For χ_1, an energetic barrier of 50 kJ mol^{-1} and a single energy minimum at 0° were obtained. Note that this result is in full agreement with previous ab initio calculations for triazole-functionalized deoxyuridine.[18] The relatively high barrier for χ_1 most likely hinders the rotation of the triazole moiety around the C_2–C_3 bond. In contrast, the energy profile of χ_2 reveals a much smaller barrier of 14 kJ mol^{-1} and two energetically equal minima around 0° and 180°. This barrier is low enough to allow for the rotation of the nitroxide moiety around the C–N bond at room temperature, which would explain the higher mobility of **dU̓** at room temperature as compared with the rigid **Çm**. However, the energetically favorable conformations of the N_1–C_5 bond with χ_2 values near 0° and 180° can be stabilized at lower temperatures, for example, at the freezing point of the solvent, which would result in a reduced **dU̓**–**dU̓** distance distribution width.

To explain the PELDOR-obtained distances, the structures of **2 dU̓** and **3 Çm** were modeled. First, the idealized structures of the non-modified RNA double strands **2** and **3** were generated by using the program 3DNA.[26] Then, the program MtsslWizard[27] was used to attach the models of the spin labels to these structures. The structural model of **dU̓** was based on the DFT-optimized structure described above. The dihedral angle χ_2 was set to 0° and 180° with equal probabilities. In addition, the dihedral angles χ_1 and χ_2 were allowed to deviate from their optimal values by ±20°, to account for small bond fluctuations. A model of **Çm** was approximated by the structure of **Ç** taken from literature.[28] The resulting models of **2 dU̓** and **3 Çm** are presented in Figure 8a. The inter-label distance predicted by the model of **2 dU̓** was 4.78 nm and deviates from the corresponding PELDOR-obtained value by only 0.1 nm (Figure 8b). In case of **3 Çm**, the predicted inter-labeled distance was 3.71 nm and differs from the PELDOR-derived distance by 0.6 nm (Figure 8b). Such deviation from the experimental value indicates that the used model of **Çm** is not accurate enough. One possible reason is that the methylation of the ribose 2′-OH group results in a different orientation of the cytidine nucleobase (Figure 9). Substituting the cytidine with its 2′-methylated derivative, the crystal structure of which is known,[22b] led to a decrease of the observed deviation to 0.3 nm (Figure 8b). The remaining difference may be attributed to a slight variation of the helix structure at the RNA ends induced by **Çm**. This assumption is consistent with the observed changes in the CD spectrum of **3** upon labeling with **Çm** (Figure 2b).

Taking into account that both **dU̓** and **Çm** were found to be fairly rigid in the glassy frozen solutions of the RNA duplexes, the widths of the PELDOR distance distributions are mainly determined by intrinsic motion of RNA. This implies that the relative dynamics of two sites of RNA can be accessed by the PELDOR measurements. This opportunity has been previously investigated in detail for **Ç**-labeled DNAs.[14] In that work, the width of the distance distribution between the two **Ç** labels was shown to depend on the relative position of these labels within the DNA sequence, which was subsequently interpreted in terms of a twist-stretch dynamical model of DNA. By using the present spin labels, the same investigation is, in principle, also possible for RNA but will be the subject of future work.

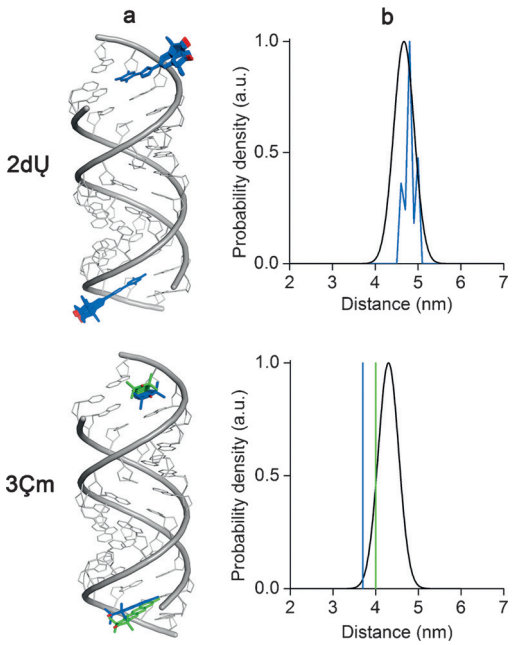


Figure 8. Structure modeling for **2 dU̓** and **3 Çm**. a) The modeled structures of the spin-labeled RNA duplexes. RNA is depicted as a gray cartoon model, and the spin labels are drawn as blue sticks with a red-colored oxygen atom. The model of **Çm** based on the crystal structure of the 2′-methylated cytidine derivative[27] is shown as green sticks. b) The corresponding inter-label distributions (blue) are overlaid with the PeldorFit-derived distance distributions (black). The green line corresponds to the inter-label distance for the green model in (a).

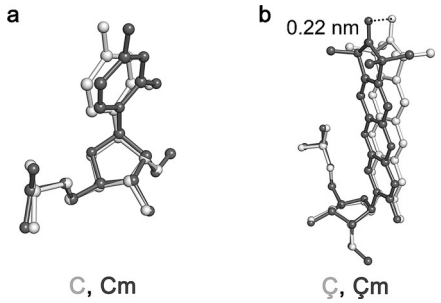


Figure 9. Effect of methylation on the structure of **Çm**. a) Overlay of the structures of cytidine **C** (light gray) and its 2′-methylated derivative **Cm** (dark gray). The structure of **C** was generated by the program 3DNA,[26] the structure of **Cm** was adopted from the crystal structure with PDB-ID 310D.[22b] b) The corresponding structures of **Ç** (light gray) and **Çm** (dark gray).

In addition, the orientation selective Q-band PELDOR data enabled the determination of the angle between the planes of the two nitroxide rings (β) and the angle between the planes of the nitroxide rings and the distance vector (90°-ξ) with a precision of 30° and 10°, respectively (Table 1). To compare these angles with the generated structural models, average values of the inclination angles were determined for the models. This yielded $\xi = 61°$ and $\beta = 56°$ for **2 dU̓**, and $\xi = 34°$ and $\beta = 24°$ for **3 Çm**. The calculated ξ values deviate by 23° and 10° from

the corresponding PeldorFit-derived angles of 38° (**2dU**) and 44° (**3Çm**). Larger deviations of 37° and 20° are found for the β angle. This difference may be attributed to the uncertainty of the PeldorFit-derived values as well as to some deviation of the modeled RNA structures from the actual structure.

Conclusions

Click chemistry on solid support was successfully used to label RNA strands with an azide-functionalized nitroxide spin label. The labeling procedure was shown to be synthetically more convenient and efficient than the analogous procedure for the spin label **Çm**. It was shown that the spin labeled nucleotide **dU** is incorporated into the RNA duplex without significantly perturbing the structure. The structure and dynamics of the obtained spin-labeled nucleotide **dU** in RNA was investigated by EPR spectroscopy and DFT calculations. Both revealed that the nitroxide moiety of **dU** is flexible to some extent at room temperature. However, lowering the temperature to the freezing temperature of the solvent seems to lock it in its energetically favorable conformation, which corresponds to an almost planar geometry of the spin-labeled nucleobase.

The performance of **dU** and **Çm** for PELDOR measurements has also been explored. Both labels provided a set of orientation-selective time traces, yielding information about the inter-label distances and the relative orientations of the labels. The obtained distances and orientations are consistent with the modeled structures of the spin-labeled RNA duplexes. During the modeling, the procedure of in silico spin labeling of oligonucleotides with **dU** and **Çm** was established and implemented in the program mtsslWizard.

Taken together, our findings reveal the potential of spin labeling through click chemistry in the structural and dynamical studies of more complex RNA structures. Moreover, this approach is not restricted to the present spin label only, but allows versatile application of a wide range of azide-functionalized nitroxides as spin labels.

Experimental Section

Synthesis and spin labeling

Compound **Çm** was synthesized by using the procedure described by Höbartner et al.[12] The RNA strands **3** and **3Çm** were purchased from Ella Biotech.

The nitroxide-functionalized azide **1** was synthesized according to the synthesis introduced by Jakobsen et al.[17] The RNA strand **2** and its 5-ethynyl-2′-dU modified counterpart on CPG support were also purchased from Ella Biotech. The latter strand was placed into a 1.5 mL reaction tube and dried under vacuum overnight. CuI (0.5 mg), THPTA (5.0 mg) and spin label **1** (5.0 mg) were placed in a small glass reaction vial with a magnetic stir bar. This vial was placed in a 10 mL Schlenk flask and dried under vacuum overnight. DMSO (100 μL), which was degassed three times by using the freeze-pump-thaw method, was then added into the reaction vial under argon atmosphere. The obtained solution was stirred for 2 min, then the reaction mixture was added into the reaction tube with RNA attached to the CPG support. The reaction tube was filled with argon and sealed by using Parafilm and then incubated in a thermomixer (350 rpm) at RT for 24 h. Subsequently, the solid support was washed with CH_3CN (3×100 μL). The RNA was cleaved from the CPG support and the amino groups were deprotected by adding 2 mL of a mixture of NH_3 (32%) and MeOH in a ratio of 3:1 (v/v) and overnight incubation at ambient temperature. The reaction mixture was filtered and the NH_3 removed in a speed vac. The TBDMS groups were cleaved with 1 M TBAF in THF over 16 h at RT. After desalting through a NAP-10 column (GE Healthcare) the RNA strands were purified by reverse-phase HPLC.

The obtained **2dU** and **3Çm** strands were dissolved in phosphate buffer (145 mM NaCl, 10 mM $Na_2HPO_4{\cdot}2H_2O$, 10 mM $Na_2H_2PO_4{\cdot}2H_2O$, pH 7.0). To ensure duplex formation, the samples were annealed by heating to 70 °C for 1 min and then cooling to 5 °C. For the CW-EPR measurements, 10 μL aliquots of 100 μM RNA solutions were prepared. For the PELDOR measurements, the RNA solutions were lyophilized and then dissolved in sterile filtered D_2O and 20% deuterated ethylene glycol, yielding final concentration of 100 μM.

CD and UV measurements

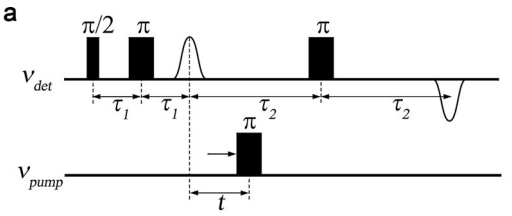
CD spectra were recorded with a Jasco J-810 spectrophotometer. The measurements were performed on 10 μM RNA samples in phosphate buffer (145 mM NaCl, 10 mM $Na_2HPO_4{\cdot}2H_2O$, 10 mM $Na_2H_2PO_4{\cdot}2H_2O$, pH 7.0). The spectra were recorded with 100 nm min^{-1} scanning speed and three averages.

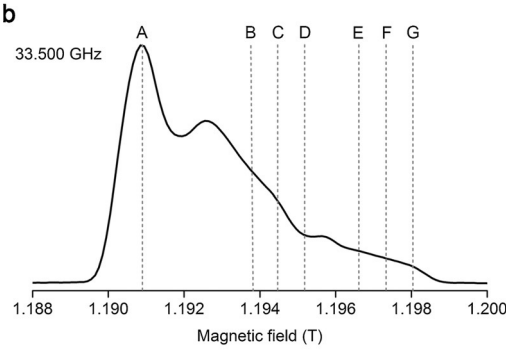
UV melting curves were recorded with a Jasco V-630 UV/Vis spectrophotometer. RNA samples (3 μM) in phosphate buffer were used. Temperature-dependent UV absorbance was measured at 260 nm, with a heating and cooling rate of 1 °C min^{-1}. Two full heating and cooling cycles were performed.

EPR measurements

The cw X-band EPR spectra were recorded with a cw X-Band EPR spectrometer EMXmicro (Bruker) equipped with a standard resonator (4119HS). All EPR spectra were collected at RT with a microwave power of 2 mW, a modulation frequency of 100 kHz, a modulation amplitude of 1 G, and a time constant of 82 ms.

PELDOR measurements were carried out with a Bruker ELEXSYS E580 spectrometer using a Flexline probe head with a Q-band resonator (Bruker, ER5106QT-2). All microwave pulses were amplified with a 150 W TWT amplifier (model 187Ka). To obtain low temperatures, a continuous flow helium cryostat (Oxford Instruments, CF935) and a temperature control system (Oxford Instruments, ITC 503S) were employed. PELDOR experiments were performed with the standard four-pulse sequence (Figure 10a). The frequencies of the pump pulse (ν_{pump}) and the detection pulses (ν_{det}) were set to be in resonance with one of the spectral positions shown in Figure 10b. The pump frequency was always set to the resonance frequency of the resonator (33.500 GHz). The chosen combinations of the pump and detection positions are listed in Figure 10c. They correspond to frequency offsets ranging from −80 to 200 MHz. The $\pi/2$ and π pulses of the detection sequence had lengths of 12 and 24 ns, respectively, and the pump pulse was 14 ns long. For the $\pi/2$ pulse a two-step phase cycle was used. The τ_1 and τ_2 intervals were set to 200 ns and 4 μs, respectively, which allowed 500 data points to be acquired with a 8 ns time increment. The PELDOR signal was recorded at 50 K with a repetition time of 1 ms. To achieve an acceptable SNR, the signal was averaged for 2 to 24 h depending on the field positions used.





c

Notation	Pump position	Detection position	Frequency offset (MHz)
$\Delta\nu_{AB}$	A	B	80
$\Delta\nu_{AC}$	A	C	100
$\Delta\nu_{AD}$	A	D	120
$\Delta\nu_{AE}$	A	E	160
$\Delta\nu_{AG}$	A	G	200
$\Delta\nu_{BA}$	B	A	-80
$\Delta\nu_{BE}$	B	E	80
$\Delta\nu_{BF}$	B	F	100
$\Delta\nu_{BG}$	B	G	120

Figure 10. a) PELDOR pulse sequence. b) Q-band spectrum of the **Çm**-labeled RNA. Different pump/detection positions for the PELDOR experiments are shown by dashed lines and labeled by letters. c) The notations of different frequency offsets corresponding to different combinations of the pump/detection positions are given in the table.

EPR simulations

The simulation of CW-EPR spectra was performed with the program MultiComponent,[23] which uses a model of microscopic ordering with macroscopic disorder (MOMD)[29] to simulate an anisotropic motion of a nitroxide. The spectra of **2**·**dU̯** and **3**·**Çm** were simulated with two components. All parameters of the simulations are given in Table S1.

The orientation-selective PELDOR time traces were analyzed with the program PeldorFit.[25] The experimental PELDOR parameters, described in the previous section, and the spectroscopic parameters of the nitroxide spins (Table S1) were used as input data for the PeldorFit calculations. The parameters of the PeldorFit model (Figure 5) were varied in the ranges given in Figure 6.

Molecular modeling

DFT calculations on the nucleobase of **dU̯** were performed with the Orca[30] software using the B3LYP functional, the def2-TZVP basis set, unrestricted spin-wave functions, and D3 dispersion corrections. The vibrational frequencies for the geometry-optimized structure were all positive, indicating that the structure represents an energy minimum. Torsional energy profiles for the dihedral angles χ_1 and χ_2 of **dU̯** (Figure 7) were obtained by using a relaxed scan.

The structural models of the RNA duplexes **2** and **3** were obtained with the program 3DNA[26] using the mode "fiber". To introduce **dU̯** or **Çm** into these models, the program mtsslWizard[27] was employed. The DFT optimized geometry of **dU̯** was used by mtsslWizard as a starting structure of the spin-labeled nucleotide, which was aligned to the uracil base of the initial RNA structure during the in silico labeling. Additionally, the dihedral angles χ_1 and χ_2 of **dU̯** (Figure 7) were allowed to deviate by $\pm 20^\circ$ from their optimized values of 0° and $0^\circ/180^\circ$, respectively.

Acknowledgements

Funding from the DFG within the SFB813 "Chemistry at Spin Centers" is gratefully acknowledged. D. Abdullin thanks Dzmitry Firaha for help with DFT calculations.

Keywords: conformation analysis · EPR spectroscopy · oligonucleotides · RNA structures · spin label

[1] a) A. D. Milov, K. M. Salikhov, M. D. Shirov, *Fiz. Tverd. Tela* **1981**, *23*, 975–982; b) R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H. W. Spiess, *Angew. Chem. Int. Ed.* **1998**, *37*, 2833–2837; *Angew. Chem.* **1998**, *110*, 2993–2998.

[2] a) G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446; b) O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1–53; c) H. El Mkami, D. G. Norman, *Methods Enzymol.* **2015**, *564*, 125–152.

[3] C. Altenbach, T. Marti, H. Khorana, W. Hubbell, *Science* **1990**, *248*, 1088–1092.

[4] a) A. Spaltenstein, B. H. Robinson, P. B. Hopkins, *J. Am. Chem. Soc.* **1988**, *110*, 1299–1301; b) W. Bannwarth, D. Schmidt, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 977–980; c) O. Schiemann, N. Piton, J. Plackmeyer, B. E. Bode, T. F. Prisner, J. W. Engels, *Nat. Protoc.* **2007**, *2*, 904–923; d) O. Schiemann, N. Piton, Y. Mu, G. Stock, J. W. Engels, T. F. Prisner, *J. Am. Chem. Soc.* **2004**, *126*, 5722–5729; e) N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, J. W. Engels, *Nucleic Acids Res.* **2007**, *35*, 3128–3143.

[5] a) A. M. MacMillan, G. L. Verdine, *J. Org. Chem.* **1990**, *55*, 5931–5933; b) A. Ramos, G. Varani, *J. Am. Chem. Soc.* **1998**, *120*, 10992–10993; c) O. Duss, M. Yulikov, G. Jeschke, F. H.-T. Allain, *Nat. Commun.* **2014**, *5*, 1–9; d) O. Duss, E. Michel, M. Yulikov, M. Schubert, G. Jeschke, F. H.-T. Allain, *Nature* **2014**, *509*, 588–592.

[6] a) T. E. Edwards, S. T. Sigurdsson, *Nat. Protoc.* **2007**, *2*, 1954–1962; b) S. Saha, A. P. Jagtap, S. T. Sigurdsson, *Methods Enzymol.* **2015**, *563*, 397–414.

[7] a) P. Z. Qin, S. E. Butcher, J. Feigon, W. L. Hubbell, *Biochemistry* **2001**, *40*, 6929–6936; b) G. P. G. Grant, N. Boyd, D. Herschlag, P. Z. Qin, *J. Am. Chem. Soc.* **2009**, *131*, 3136–3137.

[8] a) A. Okamoto, T. Inasaki, I. Saito, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3415–3418; b) G. Sicoli, F. Wachowius, M. Bennati, C. Höbartner, *Angew. Chem. Int. Ed.* **2010**, *49*, 6443–6447; *Angew. Chem.* **2010**, *122*, 6588–6592.

[9] P. Z. Qin, K. Hideg, J. Feigon, W. L. Hubbell, *Biochemistry* **2003**, *42*, 6772–6783.

[10] E. S. Babaylova, A. V. Ivanov, A. A. Malygin, M. A. Vorobjeva, A. G. Venyaminova, Y. F. Polienko, I. A. Kirilyuk, O. A. Krumkacheva, M. V. Fedin, G. G. Karpova, E. G. Bagryanskaya, *Org. Biomol. Chem.* **2014**, *12*, 3129–3136.

[11] N. Barhate, P. Cekan, A. P. Massey, S. T. Sigurdsson, *Angew. Chem. Int. Ed.* **2007**, *46*, 2655–2658; *Angew. Chem.* **2007**, *119*, 2709–2712.

[12] C. Höbartner, G. Sicoli, F. Wachowius, D. B. Gophane, S. T. Sigurdsson, *J. Org. Chem.* **2012**, *77*, 7749–7754.

[13] a) O. Schiemann, P. Cekan, D. Margraf, T. F. Prisner, S. T. Sigurdsson, *Angew. Chem. Int. Ed.* **2009**, *48*, 3292–3295; *Angew. Chem.* **2009**, *121*, 3342–3345; b) I. Tkach, S. Pornsuwan, C. Höbartner, F. Wachowius, S. T. Sigurdsson, T. Y. Baranova, U. Diederichsen, G. Sicoli, M. Bennati, *Phys. Chem. Chem. Phys.* **2013**, *15*, 3433–3437.

[14] a) A. Marko, V. Denysenkov, D. Margraf, P. Cekan, O. Schiemann, S. T. Sigurdsson, T. F. Prisner, *J. Am. Chem. Soc.* **2011**, *133*, 13375–13379; b) T. F. Prisner, A. Marko, S. T. Sigurdsson, *J. Magn. Reson.* **2015**, *252*, 187–198.

[15] A. H. El-Sagheer, T. Brown, *Chem. Soc. Rev.* **2010**, *39*, 1388 – 1405.
[16] P. Ding, D. Wunnicke, H. J. Steinhoff, F. Seela, *Chem. Eur. J.* **2010**, *16*, 14385 – 14396.
[17] U. Jakobsen, S. A. Shelke, S. Vogel, S. T. Sigurdsson, *J. Am. Chem. Soc.* **2010**, *132*, 10424 – 10428.
[18] P. Kočalka, N. K. Andersen, F. Jensen, P. Nielsen, *ChemBioChem* **2007**, *8*, 2106 – 2116.
[19] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596 – 2599; *Angew. Chem.* **2002**, *114*, 2708 – 2711.
[20] V. O. Rodionov, S. I. Presolski, D. D. Diaz, V. V. Fokin, M. G. Finn, *J. Am. Chem. Soc.* **2007**, *129*, 12705 – 12712.
[21] H. T. Steely, D. M. Gray, D. Lang, M. F. Maestre, *Biopolymers* **1986**, *25*, 91 – 117.
[22] a) M. Egli, P. S. Pallan, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 281 – 305; b) D. A. Adamiak, J. Milecki, M. Popenda, R. W. Adamiak, Z. Dauter, W. R. Rypniewski, *Nucleic Acids Res.* **1997**, *25*, 4599 – 4607.
[23] C. J. López, Z. Yang, C. Altenbach, W. L. Hubbell, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4306 – E4315.
[24] Z. J. Tan, S. J. Chen, *Biophys. J.* **2006**, *90*, 1175 – 1190.
[25] D. Abdullin, G. Hagelueken, R. I. Hunter, G. M. Smith, O. Schiemann, *Mol. Phys.* **2015**, *113*, 544 – 560.
[26] X.-J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108 – 5121.
[27] a) G. Hagelueken, R. Ward, J. Naismith, O. Schiemann, *Appl. Magn. Reson.* **2012**, *42*, 377 – 391; b) G. Hagelueken, D. Abdullin, R. Ward, O. Schiemann, *Mol. Phys.* **2013**, *111*, 2757 – 2766; c) G. Hagelueken, D. Abdullin, O. Schiemann, *Methods Enzymol.* **2015**, *563*, 595 – 622.
[28] T. E. Edwards, P. Cekan, G. W. Reginsson, S. A. Shelke, A. R. Ferré-D'Amaré, O. Schiemann, S. T. Sigurdsson, *Nucleic Acids Res.* **2011**, *39*, 4419 – 4426.
[29] D. E. Budil, S. Lee, S. Saxena, J. H. Freed, *J. Magn. Reson.* **1996**, *120*, 155 – 189.
[30] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73 – 78.

Received: May 19, 2016
Published online on July 14, 2016

CHEMISTRY
A European Journal

Supporting Information

Post-synthetic Spin-Labeling of RNA through Click Chemistry for PELDOR Measurements**

Mark Kerzhner[+],[a] Dinar Abdullin[+],[b] Jennifer Więcek,[b] Hideto Matsuoka,[b] Gregor Hagelueken,[b] Olav Schiemann,*[b] and Michael Famulok*[a, c]

chem_201601897_sm_miscellaneous_information.pdf

Contents

HPLC profile of the click reaction

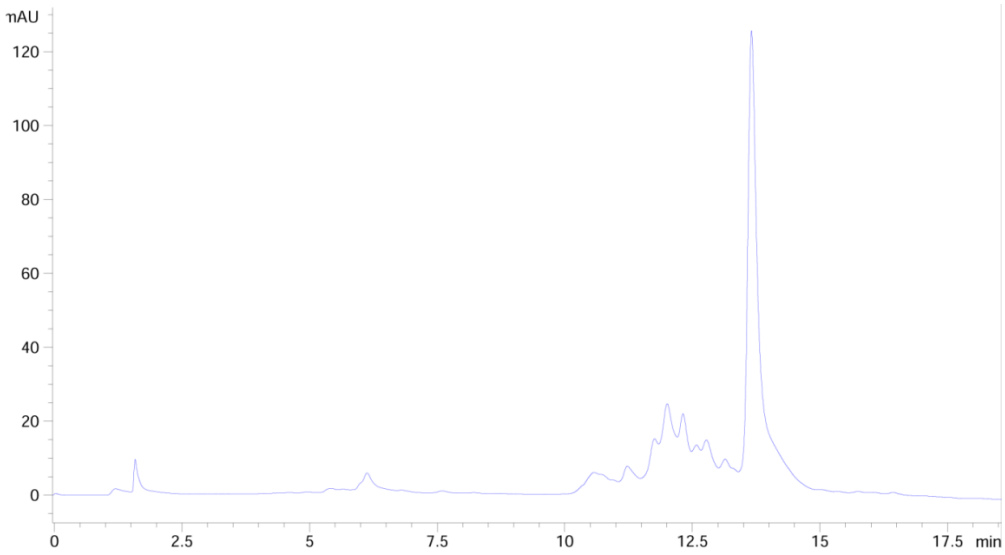
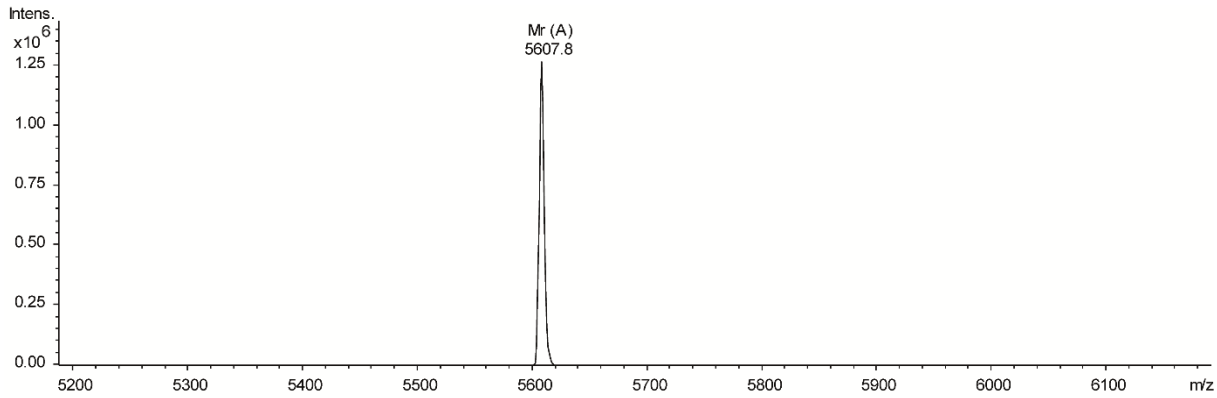


Figure S1. Crude HPLC of spin labeling on solid support for **2dU**. The most intense peak corresponds to the product.

LC-MS analysis

a



b

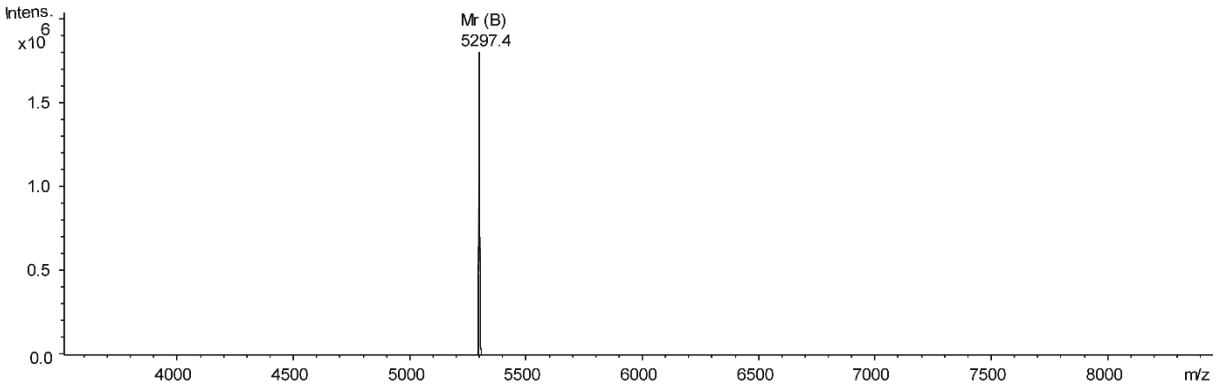
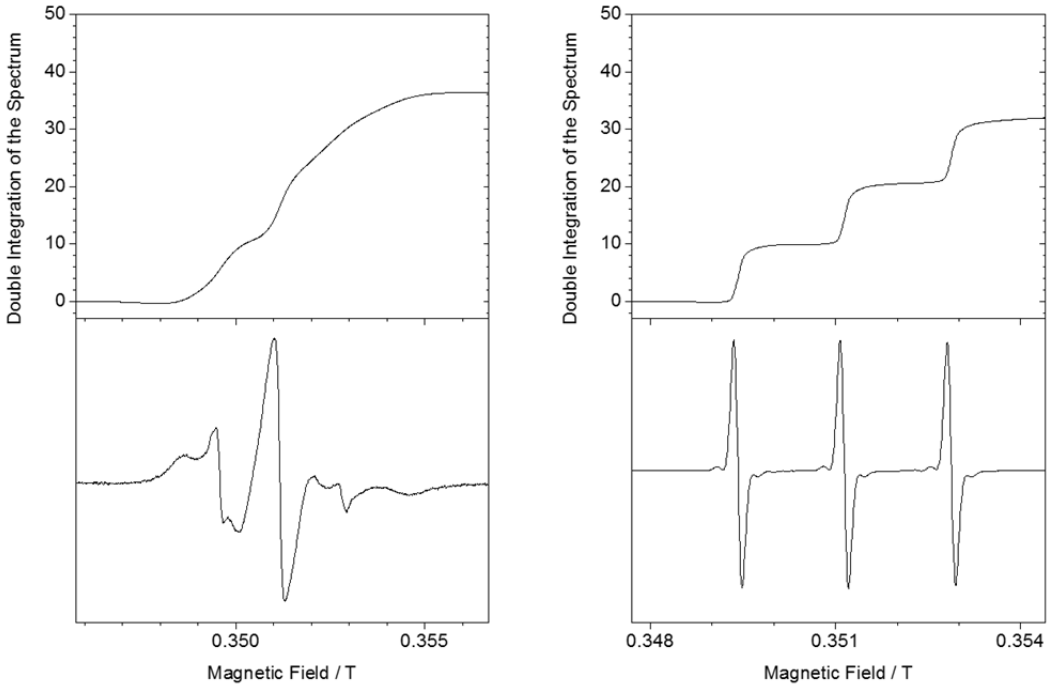


Figure S2. ESI Negative Mode of a) **2dU**, calcd. 5607.8 m/z (DMT-on), found 5607.8 m/z (DMT-on), b) **3Ç**m, calcd. 5297.4 m/z, found 5297.4 m/z.

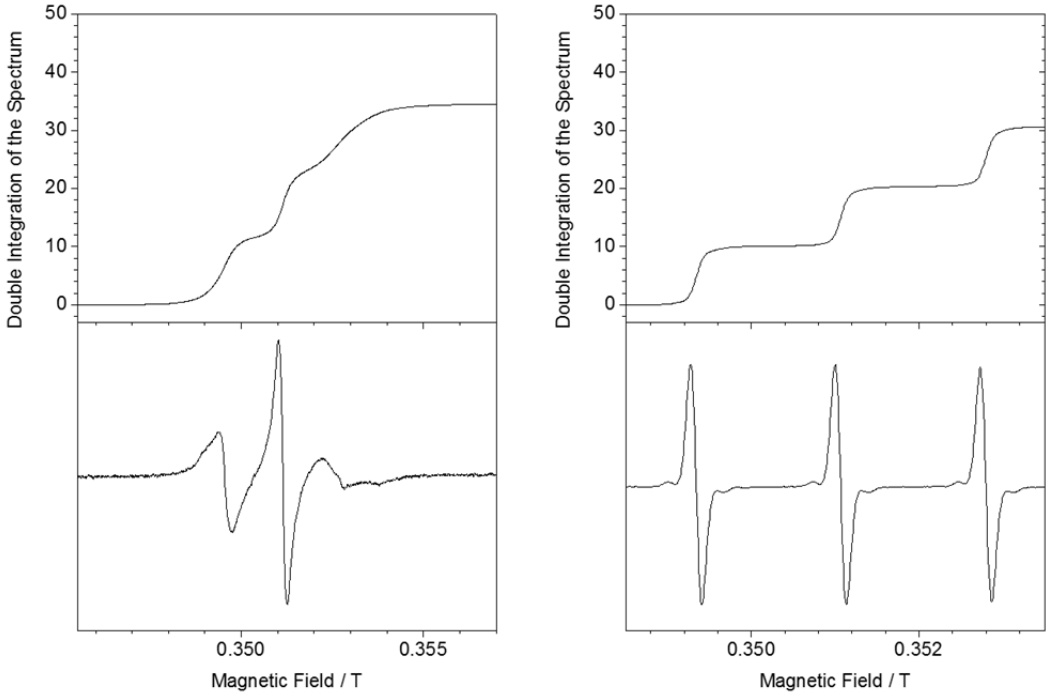
Determination of Spin Labeling Efficiency of 3Çm

Concentration of the RNA duplex: 50 μM
Calculated spin concentration: 100 μM (TEMPO) x (36.4 / 32.1) = 113 μM (**Çm**)



Determination of Spin Labeling Efficiency of 2dỤ

Concentration of the RNA duplex: 50 μM
Calculated spin concentration: 100 μM (TEMPO) x (36.5 / 30.7) = 112 μM (**dỤ**)



Thermal denaturation

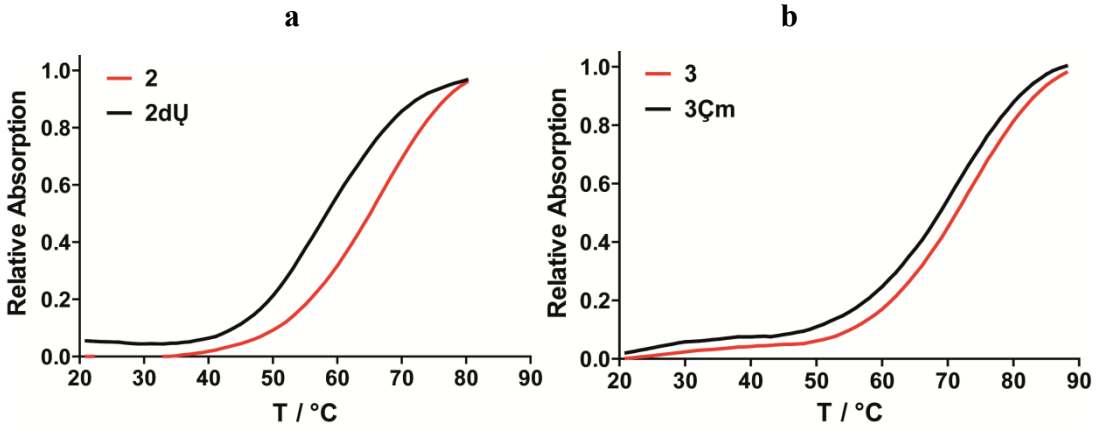


Figure S3. a) Melting curves of **2** (T_m = 66°C) and **2dU** (T_m = 58°C). **b)** Melting curves of **3** (T_m = 73°C) and **3Çm** (T_m = 72°C).

cw-EPR spectra simulations

Table S1. Parameters used for the simulation of the CW-EPR spectra of **2dU** and **3Çm** with the program MultiComponent.

Parameter	2dU		3Çm	
	slow-motion fraction	fast-motion fraction	slow-motion fraction	fast-motion fraction
g[a]	[2.0086, 2.0064, 2.0026]		[2.0086, 2.0064, 2.0026]	
$A(^{14}N)$ (mT)[b]	[0.58, 0.58, 3.67]		[0.58, 0.58, 3.67]	
log(R)[c]	[7.97, 1.66]	8.27	[7.32, 1.26]	8.31
τ_c (ns)[d]	1.8	0.9	7.9	0.8
S_{20}[e]	0.13	-	0.70	-
α_D, β_D, γ_D[f]	0°, 31°, 0°	-	0°, 42°, 0°	-
ΔH (mT)[g]	0.0		0.2	
Weight (%)	0.93	0.07	0.93	0.07

[a] The g-factor is given in a form [g_{xx}, g_{yy}, g_{zz}].
[b] The hyperfine coupling tensor is given in a form [A_{xx}, A_{yy}, A_{zz}].
[c] The logarithm of diffusion tensor R (s^{-1}): axial for the slow-motion fraction of spin labels and isotropic for fast-motion fraction of spin labels.
[d] The effective rotational correlation time.
[e] The order parameter.
[f] A set of Euler angles describing the relative orientation of the diffusion tensor axes with respect to the nitroxide magnetic frame.
[g] The inhomogeneous line broadening.

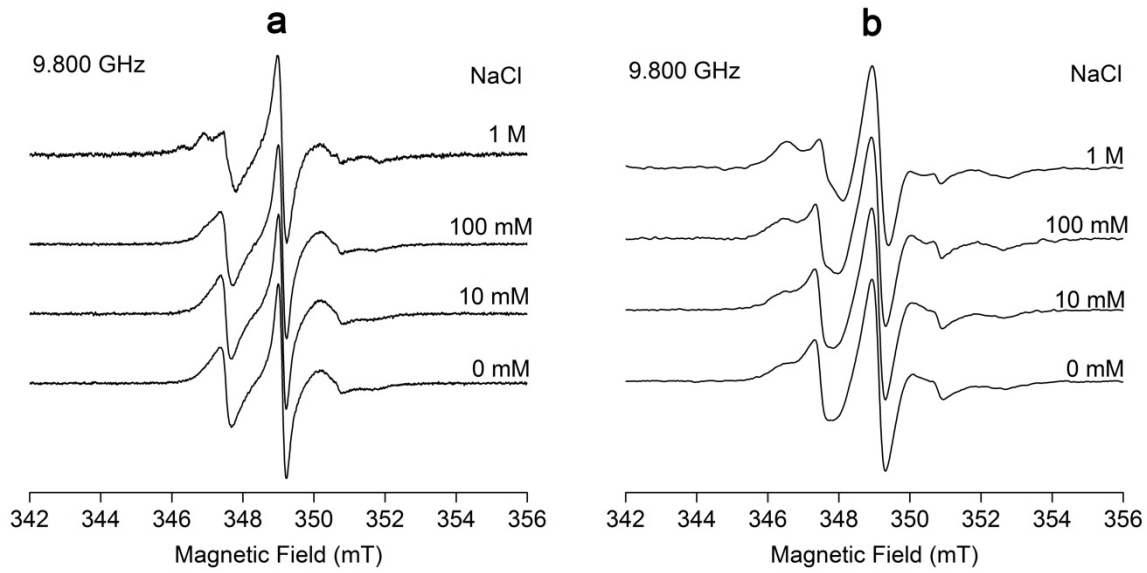


Figure S4. Dependence of the EPR spectrum of **a) 2dU̧** and **b) 3Çm** on the NaCl concentration in the samples.

Table S2. The amount of the fast-motion fraction of the spin labels in **2dU̧** and **3Çm** for the different concentrations of NaCl in the samples.

NaCl concentration (mM)	**Fast-motion fraction (%)**	
	2dU̧	**3Çm**
0	0.14 ± 0.04	0.14 ± 0.02
10	0.12 ± 0.04	0.12 ± 0.02
100	0.08 ± 0.04	0.08 ± 0.02
145	0.07 ± 0.04	0.07 ± 0.02
1000	0.05 ± 0.04	0.05 ± 0.02

PELDOR data of 2dU and 3Cm

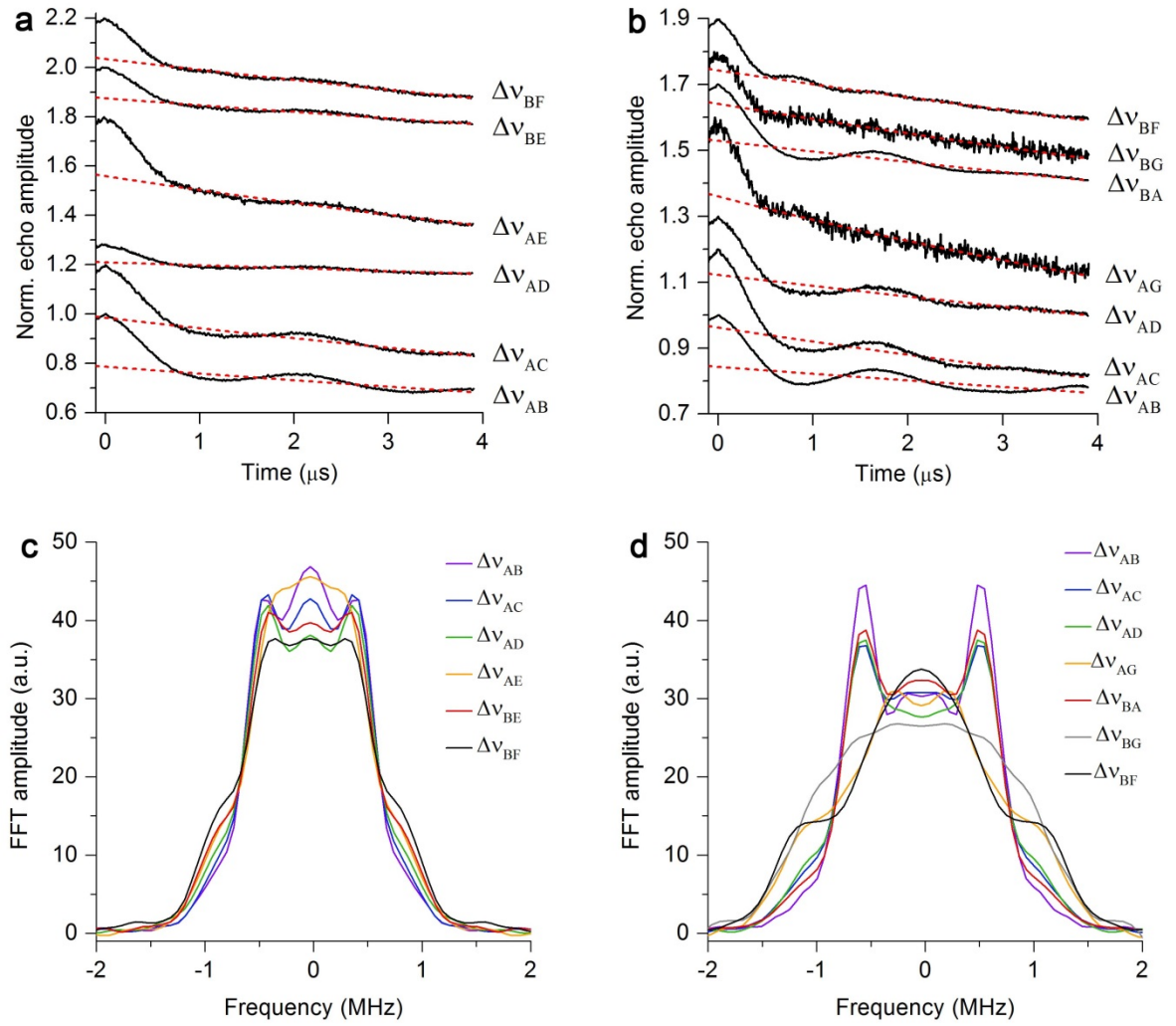


Figure S5. PELDOR time traces of **a) 2dU** and **b) 3Cm** (black lines) are overlaid with the background fits (red dashes). The designation of the frequency offsets is given in accordance with Fig. 10 in the main test. The Fourier transforms of the PELDOR time traces of **2dU** and **3Cm** are shown in **c)** and **d)**, respectively.

PeldorFit models of 2dU and 3Cm

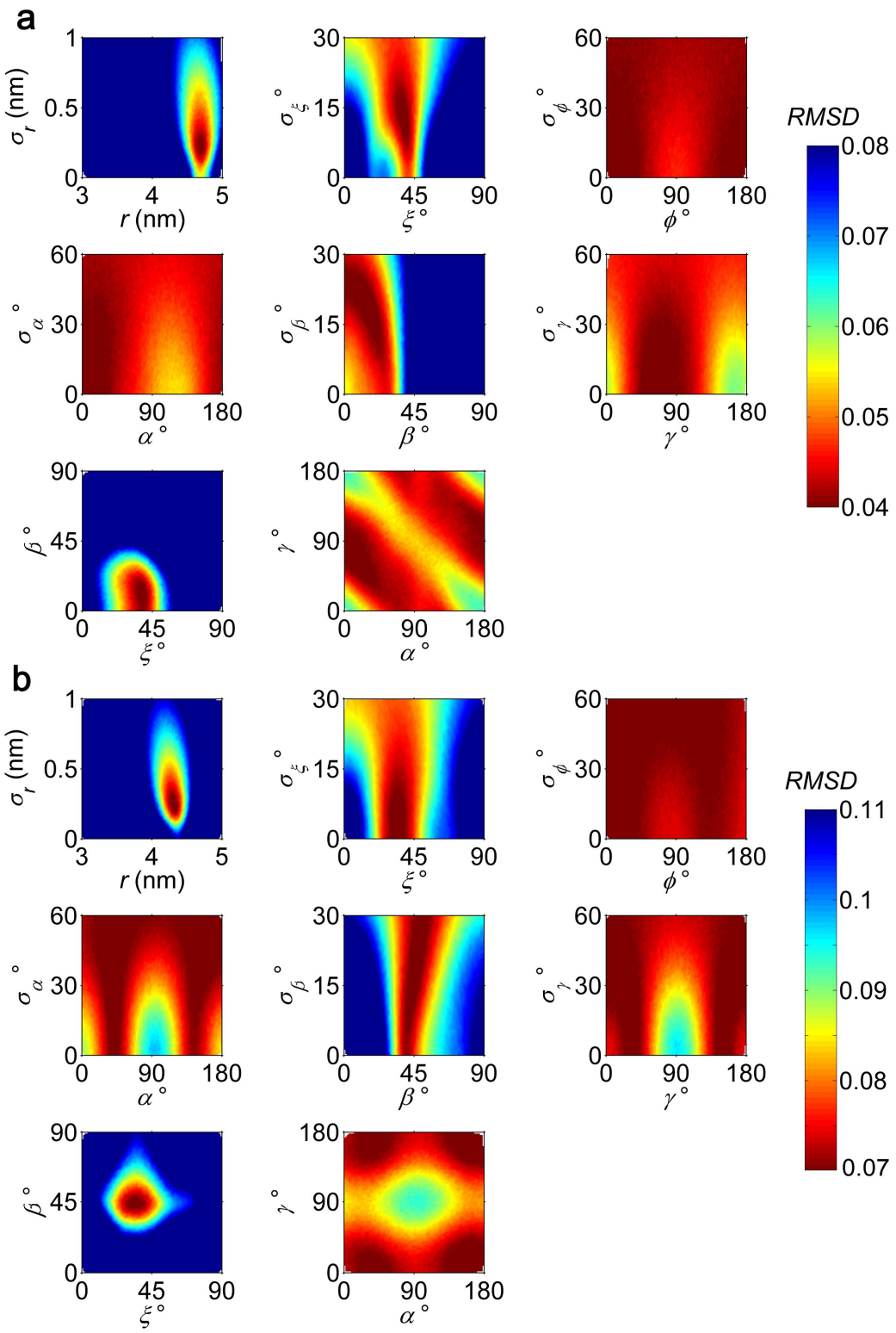


Figure S6. RMSD between experimental and simulated PELDOR time traces as a function of the geometric parameters of PeldorFit model for **a) 2dU** and **b) 3Cm**. The lowest values of the RMSD scales, which are shown by dark red, correspond to the 110% of the minimal RMSDs.

Table S3. Summary of the symmetry-related sets of angular parameters for **2dU**. All sets of angles correspond to the RMSD value of 0.037±0.001.

Transformation[a]	ξ (°)	φ (°)	α (°)	β (°)	γ (°)
Fitting result	38	8	19	11	66
180° rotation about g_{xx}^{A}	142	352	161	169	246
180° rotation about g_{yy}^{A}	142	172	341	169	246
180° rotation about g_{zz}^{A}	38	188	199	11	66
180° rotation about g_{xx}^{B}	38	8	199	169	114
180° rotation about g_{yy}^{B}	38	8	199	169	294
180° rotation about g_{zz}^{B}	38	8	19	11	246
180° rotation about g_{xx}^{A} and g_{xx}^{B}	142	352	341	11	294
180° rotation about g_{xx}^{A} and g_{yy}^{B}	142	352	341	11	114
180° rotation about g_{xx}^{A} and g_{zz}^{B}	142	352	161	169	66
180° rotation about g_{yy}^{A} and g_{xx}^{B}	142	172	161	11	294
180° rotation about g_{yy}^{A} and g_{yy}^{B}	142	172	161	11	114
180° rotation about g_{yy}^{A} and g_{zz}^{B}	142	172	341	169	66
180° rotation about g_{zz}^{A} and g_{xx}^{B}	38	188	19	169	114
180° rotation about g_{zz}^{A} and g_{yy}^{B}	38	188	19	169	294
180° rotation about g_{zz}^{A} and g_{zz}^{B}	38	188	199	11	246

[a] g_{xx}^{A}, g_{yy}^{A}, and g_{zz}^{A} denote the principal components of the *g*-tensor of spin A; g_{xx}^{B}, g_{yy}^{B}, and g_{zz}^{B} denote the *g*-tensor of spin B.

Table S4. Summary of the symmetry-related sets of angular parameters for **3Çm**. All sets of angles correspond to the RMSD value of 0.067±0.001.

Transformation[a]	ξ (°)	φ (°)	α (°)	β (°)	γ (°)
Fitting result	34	17	159	44	174
180° rotation about g_{xx}^{A}	146	343	21	136	354
180° rotation about g_{yy}^{A}	146	163	201	136	354
180° rotation about g_{zz}^{A}	34	197	339	44	174
180° rotation about g_{xx}^{B}	34	17	339	136	6
180° rotation about g_{yy}^{B}	34	17	339	136	186
180° rotation about g_{zz}^{B}	34	17	159	44	354
180° rotation about g_{xx}^{A} and g_{xx}^{B}	146	343	201	44	186
180° rotation about g_{xx}^{A} and g_{yy}^{B}	146	343	201	44	6
180° rotation about g_{xx}^{A} and g_{zz}^{B}	146	343	21	136	174
180° rotation about g_{yy}^{A} and g_{xx}^{B}	146	163	21	44	186
180° rotation about g_{yy}^{A} and g_{yy}^{B}	146	163	21	44	6
180° rotation about g_{yy}^{A} and g_{zz}^{B}	146	163	201	136	174
180° rotation about g_{zz}^{A} and g_{xx}^{B}	34	197	159	136	6
180° rotation about g_{zz}^{A} and g_{yy}^{B}	34	197	159	136	186
180° rotation about g_{zz}^{A} and g_{zz}^{B}	34	197	339	44	354

[a] g_{xx}^{A}, g_{yy}^{A}, and g_{zz}^{A} denote the principal components of the *g*-tensor of spin A; g_{xx}^{B}, g_{yy}^{B}, and g_{zz}^{B} denote the *g*-tensor of spin B.

Publications

Parts of this thesis have been published in peer-reviewed journals:

[P1] D. Abdullin, G. Hagelueken, R. I. Hunter, G.M. Smith, O. Schiemann, "Geometric model-based fitting algorithm for orientation-selective PELDOR data", *Mol. Phys.* **2015**, *113*, 544–560.

[P2] G. Hagelueken, D. Abdullin, R. Ward, O. Schiemann, "mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL, *Mol. Phys.* **2013**, *111*, 2757–2766."

[P3] G. Hagelueken, D. Abdullin, O. Schiemann, "mtsslSuite: Probing biomolecular conformation by spin labeling studies", *Meth. Enzym.* **2015**, *563*, 595–622.

[P4] D. Abdullin, N. Florin, G. Hagelueken, O. Schiemann, "EPR-Based Approach for the Localization of Paramagnetic Metal Ions in Biomolecules", *Angew. Chem. Int. Ed.* **2015**, 54, 1827–1831.

[P5] D. Abdullin, G. Hagelueken, O. Schiemann, "Determination of Nitroxide Spin Label Conformations *via* PELDOR and X-ray Crystallography", *Phys. Chem. Chem. Phys.* **2016**, *18*, 10428–10437.

[P6] D. Abdullin, F. Duthie, A. Meyer, E.S. Müller, G. Hagelueken, O. Schiemann, Comparison of PELDOR and RIDME for Distance Measurements between Nitroxides and Low-Spin Fe(III) Ions, *J. Phys. Chem. B* **2015**, *119*, 13534–13542.

[P7] M. Kerzhner, D. Abdullin, J. Więcek, H. Matsuoka, G. Hagelueken, O. Schiemann, M. Famulok, "Post-synthetic Spin Labeling of RNA via Click Chemistry for PELDOR Measurements", *Chem. Eur. J.* **2016**, *22*, 12113–12121.

Further publications:

[P8] S.C. Roy, R. Glaum, D. Abdullin, O. Schiemann, N. Quang Bac, K.-H. Lii, "Solid Solution Formation between Vanadium(V) and Tungsten(V) Oxide Phosphate", *Z. Anorg. Allg. Chem.* **2014**, *640*, 1876–1885.

[P9] A. Meyer, D. Abdullin, G. Schnakenburg, O. Schiemann, Single and Double Nitroxide Labeled Bis(terpyridine)copper(II): The Influence of Multispin and Jahn-Teller Effects on PELDOR and RIDME, *Phys. Chem. Chem. Phys.* **2016**, *18*, 9262–9271.

Conference presentations:

[P10] Poster: "PELDOR based Triangulation of Metal Ions in Biomacromolecules", *Workshop "Pulse Dipolar ESR Spectroscopy (PDS) and Biomacromolecular Structure and Function: Instrumentation, Methods and Applications"*, **2013**, Cornell University, USA.

[P11] Poster: "PELDOR based Triangulation of Metal Ions in Biomacromolecules", *35th FGMR Discussion Meeting*, **2013**, Frauenchiemsee, Germany.

[P12] Poster: "EPR based Trilateration of Paramagnetic Metal Ions in Biomacromolecules", *EUROMAR 2014 Conference*, **2014**, Zürich, Switzerland.

[P13] Poster: "Comparison of different EPR techniques for Fe(III)-nitroxide distance measurements", *The 48th Annual International Meeting of the ESR Spectroscopy Group of the Royal Society of Chemistry*, **2015**, Southampton, UK.

[P14] Talk: "Trilateration of Paramagnetic Metal Ions in Biomolecules", *3rd Awaji International Workshop on "Electron Spin Science & Technology: Biological and Materials Science Oriented Applications"*, **2015**, Awaji Island, Japan.

[P15] Talk: "An EPR-based molecular GPS", *GDCh-Vortrag "Junge Forscher der Bonner Chemie stellen sich vor"*, **2015**, Bonn, Germany.

[P16] Poster: "Comparison of different EPR techniques for Fe(III)-nitroxide distance measurements", *EUROMAR 2015 Conference*, 2015, Prague, Czech Republic.

[P17] Talk: "Trilateration of Paramagnetic Metal Ions in Biomolecules", *37th FGMR Discussion Meeting*, **2015**, Darmstadt, Germany.

[P18] Poster: "EPR-based Localization of Metal Ions in Biomolecules", *10th Conference of the European Federation of EPR groups*, **2016**, Torino, Italy.

[P19] Poster: "EPR-based Localization of Metal Ions in Biomolecules", *38th FGMR Discussion Meeting*, **2016**, Düsseldorf, Germany.

Abbreviations

CD	Circular dichroism
CW-EPR	Continuous wave electron paramagnetic resonance
DEER	Double electron-electron resonance
DFT	Density functional theory
DQC	Double quantum coherence
ENDOR	Electron Nuclear Double Resonance
EPR	Electron paramagnetic resonance
ESEEM	Electron spin echo envelope modulation
FFT	Fast Fourier transform
FRET	Fluorescence resonance energy transfer
HYSCORE	Hyperfine sublevel correlation spectroscopy
MD	Molecular dynamics
MTSSL	(1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate spin label
NMR	Nuclear magnetic resonance
OD	Optical density
PDB	Protein data bank
PELDOR	Pulsed electron-electron paramagnetic resonance
RIDME	Relaxation-induced dipolar modulation enhancement
RMSD	Root-mean-square deviation
SDSL	Site-directed spin labeling
SIFTER	Single frequency technique for refocusing dipolar couplings
SNR	Signal-to-noise ratio